Macroeconomía I - (Primera parte)
Apunte de cátedra

Otoño 2024
Profesor: Eduardo Engel

Jaime Aránguiz R-T. [1]

Última actualización: 25 de diciembre de 2024.

[1] Este resumen está hecho en base a las clases y diapositivas del profesor Eduardo Engel, para la cátedra de Macroeconomía I del Magíster en Economía. Para cualquier corrección o comentario favor escribir a jaranguizr@fen.uchile.cl.

Índice

SERIES DE TIEMPO

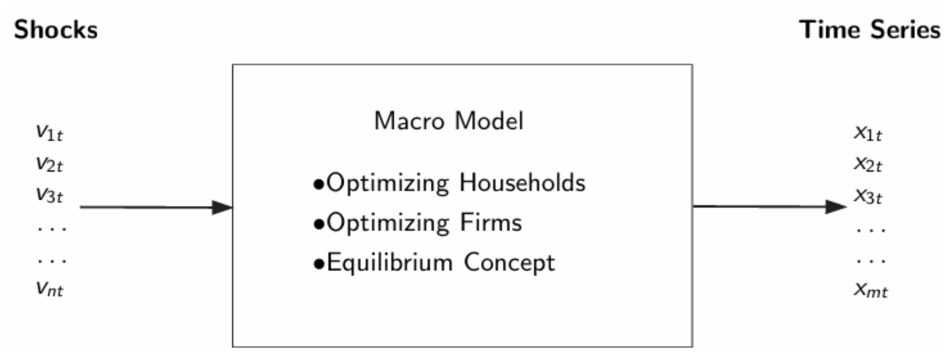
1. Series de Tiempo: Conceptos Teóricos

1.1. Macro de posgrado

Al igual que en pregrado nos van a interesar las variables agregadas (producto, inversión, consumo, empleo, etc. Sin embargo, ahora tomaremos más en serio la dinámica de estas variables, dejando de lado supuestos como el agente representativo. Jugarán un rol:

- Las expectativas.
- Los Costos de ajuste.
- La capacidad de compromiso de los agentes.

En general, todo modelo macroeconómico se puede resumir al siguiente diagrama:



Es decir, los shocks (e.g. TPM), que asumimos exógenos, definen las condiciones iniciales del modelo, que implica algún equilibrio, lo que da como resultado series de tiempo de nuestras variables de interés (e.g. producto).

Si tenemos n shocks exógenos y m variables endógenas, entonces usualmente necesitaremos que $n \leq m$, para poder identificar los parámetros. A los shocks también los llamaremos **innovaciones**, pues son el componente "nuevo" del periodo t. Definir qué variables son exógenas queda a criterio de el/la macroeconomista.

1.2. Conceptos básicos

En este curso entenderemos las series de tiempo como un **vector de realizaciones de una variable aleatoria** (VA):

$$[x_1, x_2, \ldots, x_T]$$

Notemos que dado este set de T datos, tendremos como parámetros a estimar T esperanzas $(E[x_1], \ldots, E[x_T])$, T varianzas $(Var[x_1], \ldots, Var[x_T])$ y $\frac{1}{2}T(T-1)$ correlaciones[2] $\rho(x_i, x_j)$. Es decir, tenemos un **problema**: hay muchos más parámetros que datos. La **solución** a este problema es la **estacionariedad**, que nos reducirá la cantidad de parámetros considerablemente.

[2] Resultado de la combinatoria de elegir una pareja de dos x_t.

1.3. Estacionariedad

Existen dos tipos de estacionariedad: la débil y la fuerte. Veremos la diferencia

1.3.1. Débil

Un proceso x_t en tiempo discreto es débilmente estacionario si se cumple que:

- La esperanza de x_t no depende de t
$$E[x_t] = \mu, \quad \forall t$$
- La correlación entre x_t y x_{t+h} no depende de t para ningún h. Es decir, solo depende de la “distancia entre x_t y x_{t+h}:
$$E[x_t x_{t+h}] = \gamma_k, \quad h = 0, 1, 2, ...$$

1.3.2. Fuerte

Un proceso x_t en tiempo discreto es fuertemente estacionario si para todos los enteros positivos k, $t_1, ..., t_k$, y h la distribución conjunta de

$$(x_{t_1}, x_{t_2}, ..., x_{t_k}) \quad \text{y} \quad (x_{t_1+h}, x_{t_2+h}, ..., x_{t_k+h})$$

son iguales. Esto es, la distribución del proceso es invariante bajo cambios en el tiempo. Notemos que este supuesto de la distribución implica también los supuestos de la estacionariedad débil. Es decir, bajo estacionariedad fuerte, $E[x_t]$ no depende de t y la covarianza solo depende de la distancia entre dos x_t y x_{t+h}, es decir, h.

1.4. Modelos ARMA

1.4.1. Ruido blanco

Diremos que las innovaciones de los procesos ARMA (ε_t) son ruido blanco cuando

$$\varepsilon_t \sim i.i.d.\, N(0, \sigma_\varepsilon^2)$$

Luego relajaremos el supuesto de normalidad. Entonces, los ruidos blancos cumplen con:

1.4.2. Modelos ARMA básicos

Modelo	Formulación
AR(1)	$x_t = ax_{t-1} + \varepsilon_t$
MA(1)	$x_t = \varepsilon_t + b\varepsilon_{t-1}$
ARMA(1,1)	$x_t = ax_{t-1} + \varepsilon_t + b\varepsilon_{t-1}$
AR(p)	$x_t = \sum_{k=1}^{p} a_k x_{t-k} + \varepsilon_t$
MA(q)	$x_t = \varepsilon_t + \sum_{k=1}^{q} b_k \varepsilon_{t-k}$
ARMA(p,q)	$x_t = \sum_{k=1}^{p} a_k x_{t-k} + \varepsilon_t + \sum_{k=1}^{q} b_k \varepsilon_{t-k}$

1.5. Operador y polinomio de rezago

1.5.1. Definición

Supongamos un proceso AR(1):

$$x_t = ax_{t-1} + \varepsilon_t$$

El **operador de rezago** L, es tal que $Lx_t = x_{t-1}$. Entonces el modelo anterior lo podemos escribir como

$$x_t = aLx_t + \varepsilon_t \implies (1-aL)x_t = \varepsilon_t$$

Bajo esta formulación, el **polinomio de rezago** será el término que multiplica la variable x_t. En este caso, $a(L) \equiv 1 - aL$. En general, para un modelo ARMA(p,q) podemos definir dos polinomios de rezago:

$$a(L) = 1 - a_1L - a_2L^2 - \ldots - a_pL^p$$
$$b(L) = 1 - b_1L - b_2L^2 - \ldots - b_qL^q$$

Entonces,

$$a(L)x_t = b(L)\varepsilon_t$$

1.5.2. Representación MA(∞) de un AR(1)

Descomponiendo un proceso AR(1) podemos reescribirlo como

$$\begin{aligned}
x_t &= ax_{t-1} + \varepsilon_t \\
x_t &= a(ax_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
x_t &= a(a(ax_{t-3} + \varepsilon_{t-2}) + \varepsilon_{t-1}) + \varepsilon_t \\
&\vdots \\
x_t &= a^k x_{t-k} + (\varepsilon_t + a\varepsilon_{t-1} + a^2\varepsilon_{t-2} + \ldots + a^k\varepsilon_{t-k})
\end{aligned}$$

Pero si x_t es estacionario, $|a| < 1$ y $\lim_{k\to\infty} a^k x_{t-k} = 0$. Luego,

$$x_t = \sum_{k=0}^{\infty} a^k \varepsilon_{t-k}$$

Es decir, un MA(∞). Alternativamente, usando el operador de rezagos:

$$\begin{aligned}
(1-aL)x_t &= \varepsilon_t \\
x_t &= \frac{\varepsilon_t}{(1-aL)} \\
x_t &= \sum_{k=0}^{\infty} (aL)^k \varepsilon_t \\
x_t &= \sum_{k=0}^{\infty} a^k L^k \varepsilon_t \\
x_t &= \sum_{k=0}^{\infty} a^k \varepsilon_{t-k}
\end{aligned}$$

Donde ocupamos que $|a| < 1$ para aplicar la suma geométrica. Notemos que definiendo las raíces del polinomio de rezago $a(L) = 1 - aL$ como los z tal que

$$a(z) = 1 - az = 0$$

En el caso AR(1) tenemos que la raíz es $1/a$, donde

$$|a| < 1 \implies |1/a| > 1$$

Es decir, para poder realizar la transformación del AR(1) al MA(∞) necesitamos que el valor absoluto de la raíz del polinomio sea mayor a uno.

Podemos generalizar esta condición para el caso de un AR(p), donde necesitamos que todas las raíces de $a(z) = 1 - a_1 z - ... - a_p z^p$ tengan valor absoluto mayor a uno.

1.6. Autocovarianza y autocorrelación

1.6.1. Autocovarianza

La **autocovarianza** de un proceso estacionario x_t está definida como

$$Cov(x_t, x_{t-j}) \equiv \gamma_x(j)$$

Asumiendo un proceso de media incondicional igual a cero ($E[x_t] = 0$) y recordando la definición de covarianza[3], tenemos que

$$\gamma_x(j) \equiv E[x_t x_{t-j}]$$

Notemos que:

$$\gamma_x(0) = Var(x_t)$$
$$\gamma_x(j) = \gamma_x(-j)$$

1.6.2. Autocorrelación

La **autocorrelación** de un proceso estacionario x_t está definida como

$$Corr(x_t, x_{t-j}) \equiv \rho_x(j) = \frac{\gamma_x(j)}{\gamma_x(0)}$$

En resumen, la autocovarianza y autocorrelación entre dos x_t, x_{t-j} dependen únicamente de la distancia entre ellos (j).

1.6.3. Para un ruido blanco

Dado que los ruidos blancos se distribuyen i.i.d. $N(0, \sigma_\varepsilon^2)$, se cumplirá que

$$\gamma_\varepsilon(0) = \sigma_\varepsilon^2$$
$$\gamma_\varepsilon(j) = 0, \quad j \neq 0$$
$$\rho_\varepsilon(0) = 1$$
$$\rho_\varepsilon(j) = 0, \quad j \neq 0$$

[3] $Cov(X,Y) = E[XY] - E[X]E[Y]$.

1.6.4. Para un AR(1)

La autocovarianza será igual a:

$$\gamma_x(j) = E[x_t x_{t+j}]$$
$$\gamma_x(j) = E[x_t(ax_{t+j-1} + \varepsilon_{t+j})]$$
$$\gamma_x(j) = aE[x_t x_{t+j-1}] + E[x_t e_{t+j}]$$
$$\gamma_x(j) = a\gamma_x(j-1)$$

$$\therefore \boxed{\gamma_x(j) = a \cdot a \cdot \ldots \cdot a\gamma_x(0) = a^j \cdot Var(x_t)}$$

Entonces, la autocorrelación será

$$\rho_x(j) = \frac{Cov(x_t x_{t+j})}{DS(x_t) \cdot DS(x_{t+j})} = \frac{a^j \cdot Var(x_t)}{Var(x_t)}$$

$$\therefore \boxed{\rho_x(j) = a^j}$$

Adicionalmente, podemos mostrar que $Var(x_t) = \frac{1}{1-a^2}\sigma_\varepsilon^2$:

$$Var(x_t) = Var(ax_{t-1} + \varepsilon_t)$$
$$Var(x_t) = a^2 Var(x_{t-1}) + Var(\varepsilon_t)$$
$$Var(x_t) = a^2 Var(x_t) + \sigma_\varepsilon^2$$
$$(1-a^2)Var(x_t) = \sigma_\varepsilon^2$$
$$Var(x_t) = \frac{1}{1-a^2}\sigma_\varepsilon^2$$

1.7. Proyecciones

Para proyectar x_{t+k} basados en la información disponible en t, $\mathscr{I}_t$, es natural utilizar la media condicional:

$$E[x_{t+k}|\mathscr{I}_t] \equiv E_t[x_{t+k}]$$

Que es justamente el predictor que minimiza el error cuadrático medio de la proyección:

$$E_t[x_{t+k}] = \underset{g}{argmin}\, MSE_k = \underset{g}{argmin}\, E[(x_{t+k} - g(x_t, x_{t-1}, \ldots)^2]$$

1.7.1. Proyectando un AR(1)

Es claro ver que

$$E_t[x_{t+k}] = E_t[ax_{t+k-1} + \varepsilon_{t+k}]$$
$$E_t[x_{t+k}] = aE_t[ax_{t+k-2} + \varepsilon_{t+k-1}]$$
$$E_t[x_{t+k}] = a^2 E_t[ax_{t+k-3} + \varepsilon_{t+k-2}]$$
$$\vdots$$
$$E_t[x_{t+k}] = a^k x_t$$

Luego, para cada $k \geq 1$, el error de proyección es

$$
\begin{aligned}
x_{t+1} - E_t[x_{t+1}] &= \varepsilon_{t+1} \qquad (x_{t+1} = ax_t + \varepsilon_{t+1}) \\
x_{t+2} - E_t[x_{t+2}] &= \varepsilon_{t+2} + a\varepsilon_{t+1} \qquad (x_{t+2} = ax_{t+1} + \varepsilon_{t+2} = a(ax_t + \varepsilon_{t+1}) + \varepsilon_{t+2}) \\
x_{t+3} - E_t[x_{t+3}] &= \varepsilon_{t+3} + a\varepsilon_{t+2} + a^2\varepsilon_{t+1} \qquad (x_{t+3} = ax_{t+2} + \varepsilon_{t+3} = a(a(x_t + \varepsilon_{t+1}) + \varepsilon_{t+2}) + \varepsilon_{t+3}) \\
&\vdots \\
x_{t+k} - E_t[x_{t+k}] &= \varepsilon_{t+k} + a\varepsilon_{t+k-1} + ... + a^{k-1}\varepsilon_{t+1}
\end{aligned}
$$

Por lo tanto, el error cuadrático medio (MSE) de cada proyección será[4]:

$$
\begin{aligned}
MSE_1 &= \sigma_\varepsilon^2 \\
MSE_2 &= (1 + a^2)\sigma_\varepsilon^2 \\
&\vdots \\
MSE_k &= (1 + a^2 + \cdots + a^{2k-2})\sigma_\varepsilon^2
\end{aligned}
$$

De los resultados anteriores, concluimos que

$$
\lim_{k\to\infty} E_t[x_{t+k}] = 0, \qquad \lim_{k\to\infty} MSE_k = \frac{1}{1-a^2}\sigma_\varepsilon^2 = \sigma_x^2
$$

1.7.2. Proyectando un MA(1)

Dado el proceso MA(1):

$$
x_t = \varepsilon_t + b\varepsilon_{t-1}
$$

Tendremos la información de $\mathscr{I}_t = \{x_t, x_{t-1}, ...\}$. Además, asumimos $|b| < 1$, lo cual implica que el proceso puede ser expresado como una función de $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, ...$ tal que $\mathscr{I}_t = \{\varepsilon_t, \varepsilon_{t-1}, ...\}$. Entonces[5],

$$
\begin{aligned}
E_t[x_{t+1}] &= E_t[\varepsilon_{t+1} + b\varepsilon_t] = b\varepsilon_t \\
E_t[x_{t+2}] &= E_t[\varepsilon_{t+2} + b\varepsilon_{t-1}] = 0 \\
&\vdots \\
E_t[x_{t+k}] &= 0, \quad k = 2, 3, ...
\end{aligned}
$$

Luego, el error cuadrático medio (MSE) de cada proyección será:

$$
\begin{aligned}
MSE_1 &= \sigma_\varepsilon^2 \\
MSE_k &= E[x_{t+k}^2] = \sigma_x^2, \quad k = 2, 3, ...
\end{aligned}
$$

[4] Notar que en este caso $MSE_k = E[(x_{t+k} - E_t[x_{t+k}])^2] = E[(\varepsilon_{t+k} + a\varepsilon_{t+k-1} + ... + a^{k-1}\varepsilon_{t+1})^2]$. Luego, para $k = 2$, por ejemplo:

$$
E[(\varepsilon_{t+2} + a\varepsilon_{t+1})^2] = E[\varepsilon_{t+2}] + 2aE[\varepsilon_{t+2}\varepsilon_{t+1}] + a^2E[\varepsilon_{t+1}] = \sigma_\varepsilon^2 + 2a \cdot 0 + a^2\sigma_\varepsilon^2 = (1 + a^2)\sigma_\varepsilon^2
$$

[5] Dado que $E_t[\varepsilon_t + k] = 0$, $\forall k = 1, 2, ...$

1.8. Función impulso respuesta (IRF)

Dado un proceso x_t ARMA con innovación ε_t:

$$a(L)x_t = b(L)\varepsilon_t$$

La función impulso respuesta (IRF) de x con respecto a ε es definida como

$$IRF_k \equiv \frac{\partial x_{t+k}}{\partial \varepsilon_t}, \quad k = 0, 1, 2, ...$$

1.8.1. IRF de un AR(1)

De un proceso AR(1)

$$x_t = ax_{t-1} + \varepsilon_t$$

es directo ver que

$$IRF_0 = \frac{\partial x_t}{\partial \varepsilon_t} = \frac{\partial ax_{t-1} + \varepsilon_t}{\partial \varepsilon_t} = 1$$

y que para $k \geq 1$:

$$IRF_k = \frac{\partial x_{t+k}}{\partial \varepsilon_t} = \frac{\partial (ax_{t+k-1} + \varepsilon_{t+k})}{\partial \varepsilon_t} = a \cdot \frac{\partial x_{t+k-1}}{\partial \varepsilon_t} = a \cdot IRF_{k-1}$$

Iterando esto hasta llegar a $IRF_0 = 1$ nos da que

$$IRF_k = a^k, \quad k = 0, 1, 2, ...$$

1.8.2. IRF de un AR(p)

Considerando un proceso AR(p) de la forma

$$x_t = a_1 x_{t-1} + \cdots + a_p x_{t-p} + \varepsilon_t$$

podemos ver que

$$IRF_k = \frac{\partial x_{t+k}}{\partial \varepsilon_t} = \frac{\partial (a_1 x_{t+k-1} + \cdots + a_p x_{t+k-p} + \varepsilon_{t+k})}{\partial \varepsilon_t} = a_1 \frac{\partial x_{t+k-1}}{\partial \varepsilon_t} + \cdots + a_p \frac{\partial x_{t+k-p}}{\partial \varepsilon_t}$$

Es decir,

$$IRF_k = a_1 IRF_{k-1} + \cdots + a_p IRF_{k-p}$$

Luego, asumiendo que las p raíces del polinomio de rezago $a(z) = 1 - \sum_{k=1}^{p} a_k z^k$ son $G_1, G_2, ..., G_p$, todas diferentes, entonces:

$$IRF_k = \sum_{j=1}^{p} c_j G_j^{-k}$$

Donde las constantes c_j son determinadas por las condiciones iniciales, $IRF_0 = 1$ y $IRF_k = 0$ para $k < 0$.

1.8.3. IRF de un AR(2)

Este resultado será importante para conocer las condiciones necesarias y suficientes de una IRF "con forma de joroba". Considerando el siguiente proceso AR(2):

$$(1-a_1L)(1-a_2L)x_t = \varepsilon_t$$

Donde $a_1, a_2 \in \mathbb{R}$ tal que $0 < a_2 < a_1 < 1$. El resultado mostrará que la IRF de este proceso tendrá forma de joroba si o y solo si $a_1 + a_2 > 1$.

Demostración. Tenemos un proceso

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$$

donde $\phi_1 = a_1 + a_2$ y $\phi_2 = -a_1 a_2$. Ya vimos que

$$\begin{aligned}
IRF_0 &= 1 \\
IRF_1 &= \phi_1 IRF_0 = \phi_1 \\
IRF_k &= \phi_1 IRF_{k-1} + \phi_2 IRF_{k-2}
\end{aligned}$$

Además, podemos reescribir el proceso como

$$\begin{aligned}
x_t &= \phi_1 L x_t + \phi_2 L^2 x_t + \varepsilon_t \\
(1-\phi_1 L - \phi_2 L^2)x_t &= \varepsilon_t \\
(1-a_1L-a_2L+a_1a_2L^2)x_t &= \varepsilon_t \\
(1-a_1L)(1-a_2L)x_t &= \varepsilon_t
\end{aligned}$$

Entonces es claro ver que las raíces del polinomio AR son $1/a_1$ y $1/a_2$, y por lo tanto

$$IRF_k = C_1 a_1^k + C_2 a_2^k$$

Pero tenemos las condiciones iniciales necesarias para encontrar C_1 y C_2:

$$\begin{aligned}
IRF_0 = 1 &\implies C_1 + C_2 = 1 \\
IRF_1 = \phi_1 &\implies C_1 a_1 + C_2 a_2 = a_1 + a_2
\end{aligned}$$

Resolviendo el sistema nos queda que

$$C_1 = \frac{a_1}{a_1 - a_2}, \quad C_2 = -\frac{a_2}{a_1 - a_2}$$

Es decir, la forma general de la IRF nos queda

$$IRF_k = \frac{1}{a_1 - a_2} a_1^{k+1} - \frac{1}{a_1 - a_2} a_2^{k+1} = \frac{a_1^{k+1} - a_2^{k+1}}{a_1 - a_2}$$

Entonces, para que la IRF sea decreciente a partir de cierto k debe ocurrir que

$$IRF_k > IRF_{k+1} \iff \frac{1}{a_1 - a_2} a_1^{k+1} - \frac{1}{a_1 - a_2} a_2^{k+1} > \frac{1}{a_1 - a_2} a_1^{k+2} - \frac{1}{a_1 - a_2} a_2^{k+2}$$

Reordenando

$$\frac{1}{a_1-a_2}a_1^{k+1}-\frac{1}{a_1-a_2}a_1^{k+2}>\frac{1}{a_1-a_2}a_2^{k+1}-\frac{1}{a_1-a_2}a_2^{k+2}$$
$$\frac{1}{a_1-a_2}a_1^{k+1}(1-a_1)>\frac{1}{a_1-a_2}a_2^{k+1}(1-a_2)$$

Y, si $a_1>a_2$, podemos multiplicar esto por a_1-a_2 y nos queda

$$\left(\frac{a_1}{a_2}\right)^{k+1}=\frac{1-a_2}{1-a_1}$$

Dado que $a_1/a_2>1$, $(a_1/a_2)^{k+1}$ aumenta a medida que k es mayor, lo que existe un K para el cual la IRF es decreciente a partir de $k\geq K$.

Entonces, para que la IRF tenga forma de joroba debe ocurrir que para valores pequeños de k esta sea creciente. Tomando lo que sabemos para $k=0$ y $k=1$, la IRF es creciente en este intervalo cuando

$$IRF_0<IRF_1\iff 1<a_1+a_2$$

Lo cual concluye la prueba.

1.9. Procesos integrados

Muchas series macroeconómicas son procesos no estacionarios con tasas de crecimiento estacionarias (PIB, consumo, IPC, stock de capital, etc.). Entonces, definiremos a x_t como un **proceso integrado de orden 1** si:

- x_t es **no estacionario**.
- $\Delta x_t=x_t-x_{t-1}=(1-L)x_t$ es **estacionario**.

Entonces decimos que x_t es $I(1)$. Análogamente, para decir que x_t es un **proceso integrado de orden 2**, debe ocurrir que x_t no sea estacionario y $\Delta^2 x_t=\Delta(\Delta x_t)$ sí lo sea. En este caso, decimos que x_t es $I(2)$. Por último, un proceso x_t estacionario es referido a veces como un proceso $I(0)$.

Al mismo tiempo, cuando x_t no es estacionario y Δx_t es un ARMA(p,q) estacionario, entonces decimos que x_t es un ARIMA(p,1,q), y así con cada grado de integración.

1.9.1. Camino aleatorio

El camino aleatorio es el más importante y más simple proceso $I(1)$. Un proceso x_t sigue un camino aleatorio si existe una innovación ε_t y una constante g (drift o tendendecial) tal que:

$$x_t=g+x_{t-1}+\varepsilon_t \quad \text{o} \quad x_t=gt+x_{t-1}+\varepsilon_t$$

Tomando el primer, caso (drift) es claro que

$$\Delta x_t=g+\varepsilon_t$$

Es decir, x_t sigue un camino aleatorio si y solo si Δx_t sigue un ruido blanco (más una constante), lo que implica que x_t sigue un ARIMA(0,1,0).

1.9.2. Proyección de un camino aleatorio

Dado el siguiente camino aleatorio:

$$x_t = g + x_{t-1} + \varepsilon_t$$

Entonces tendremos que la mejor proyección es

$$\begin{aligned}
E_t[x_{t+k}] &= E_t[g + x_{t+k-1} + \varepsilon_{t+k}] \\
E_t[x_{t+k}] &= E_t[2g + x_{t+k-2} + \varepsilon_{t+k-1} + \varepsilon_{t+k}] \\
E_t[x_{t+k}] &= E_t[3g + x_{t+k-3} + \varepsilon_{t+k-2} + \varepsilon_{t+k-1} + \varepsilon_{t+k}] \\
&\vdots \\
E_t[x_{t+k}] &= kg + x_t + E_t[\varepsilon_t] + \ldots + E_t[\varepsilon_{t+k}] \\
E_t[x_{t+k}] &= kg + x_t
\end{aligned}$$

Y su varianza será

$$\begin{aligned}
Var_t[E_t[x_{t+k}]] &= Var_t[kg + x_t] \\
Var_t[E_t[x_{t+k}]] &= k\sigma_\varepsilon^2
\end{aligned}$$

Es decir, si observamos $x_1, \ldots, x_T$ para un camino aleatorio (con drift conocido), al proyectar x_{T+k}, tendremos que:

- MSE_k crece linealmente con k.
- Si x_T aumenta en una unidad, entonces $E_t[x_{T+k}]$ también aumenta en una unidad.

1.9.3. IRF de un camino aleatorio

Es directo ver que

$$\begin{aligned}
IRF_k &= \frac{\partial x_{t+k}}{\partial \varepsilon_t} \\
IRF_k &= \frac{\partial [g + x_{t+k-1} + \varepsilon_{t+k}]}{\partial \varepsilon_t} \\
&\vdots \\
IRF_k &= \frac{\partial [kg + x_t + \varepsilon_t + \ldots + \varepsilon_{t+k}]}{\partial \varepsilon_t} \\
IRF_k &= 1
\end{aligned}$$

Es decir, un shock marginal a ε_t tiene un efecto uno-a-uno en futuros valores de x_t.

2. Series de Tiempo: Aplicaciones en Macroeconomía

2.1. Descomposición ciclo-tendencia

En muchos modelos macro se enfocan en los ciclos económicos. Entonces, para testear estos modelos necesitaremos una forma de estimar el ciclo, que asumiremos que sigue un proceso estacionario. En particular, descompondremos una serie x_t en la suma de un componente de tendencia y otro componente cíclico:

$$x_t = x_t^{tr} + x_t^c \tag{2.1}$$

Notemos que bajo esta especificación bastará con estimar uno de los dos componentes, y el otro se puede despejar de (2.1).

2.1.1. Tendencia determinística

Bajo este enfoque diremos que el componente cíclico es estacionario y el componente de tendencia toma la forma de un polinomio de grado bajo. Por ejemplo, una función lineal:

$$x_t = a_0 + a_1 t + e_t \implies x_t^{tr} = a_0 + a_1 t$$

Entonces, podriamos estimar el componente de tendencia mediante OLS:

$$\hat{x}_t^{tr} = \hat{a}_0 + \hat{a}_1 t, \qquad \hat{x}_t^c = \hat{e}_t = x_t - \hat{x}_t^{tr}$$

También podríamos ajustar una tendencia cúbica, de la forma

$$x_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + e_t \implies x_t^{tr} = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

Notemos que por construcción la tendencia cúbica siempre resultará en un estimador del ciclo con menos varianza (porque estaré incluyendo más regresores en la estimación). Sin embargo, dada cierta proyección $E_t[\hat{x}_{t+k}] = \hat{a}_0 + \hat{a}_1(t+k) + \hat{\mu}(x_t^c)$, donde $\hat{\mu}(x_t^c)$ es la media del componente cíclico, vemos que el error cuadrático medio siempre estará acotado[6].

$$\lim_{k\to\infty} MSE_k \cong Var(x_t^c) < \infty$$

Gracias a que se predice la tendencia bastante bien y las proyecciones de largo plazo de un proceso estacionario son triviales (iguales a su media). Esta es una ventaja de la tendencia determinística (cuando el proceso subyacente efectivamente se comporta de esta forma). En la realidad, se ha demostrado que el error de proyección con los modelos de tendencia determinística no converge a un número acotado, sino que crece a medida que la proyección es más lejana.

2.1.2. Tendencia estocástica

En la realidad, asumir una tendencia determinística es un supuesto fuerte, con una base económica débil. Una opción más razonable es asumir que el componente de tendencia sigue un proceso integrado de orden 1 ($I(1)$) y el componente cíclico es $I(0)$. Es decir, si bien x_t no sigue un proceso estacionario, Δx_t sí lo hace. En

[6]Notemos que en este caso $MSE_k = E[(\hat{e}_t)^2] = E[(\hat{x}_t^c)^2] \cong Var(x_t^c)$.

otras palabras, la tasa de crecimiento promedio de la serie en cierto periodo es aleatoria (pero estacionaria). Por ejemplo, la serie podría seguir un camino aleatorio tal que

$$x_t = x_t^{tr} + x_t^c \implies \Delta x_t = e_t$$

Ahora, como vimos en la sección anterior, el error cuadrático medio de las proyecciones de un camino aleatorio crecerá linealmente a medida que mi proyección es más alejada. Esto implica que

$$\lim_{k\to\infty} MSE_k = \infty$$

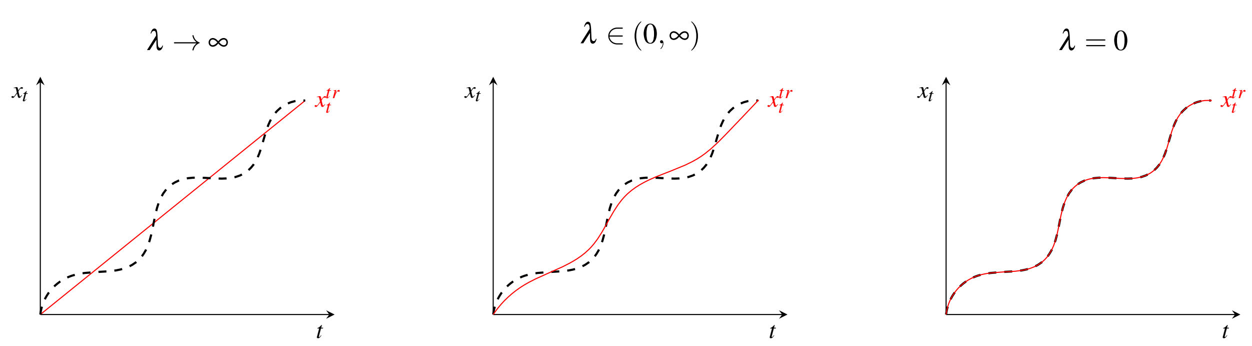
2.1.3. Filtro de Hodrick-Prescott (1980)

El filtro HP nos permite generalizar la estimación del componente de tendencia (y por tanto de ciclo) a la siguiente fórmula:

$$\min_{x_1^{tr},\ldots,x_T^{tr}} \quad \sum_{t=1}^{T}(x_t - x_t^{tr})^2 + \lambda \sum_{t=2}^{T-1}(\Delta x_{t+1}^{tr} - \Delta x_t^{tr})^2$$

En palabras, el filtro HP encuentra la trayectoria tendencial $(x_1^{tr},\ldots,x_T^{tr})$ que minimiza la suma de desviaciones al cuadrado de la tendencia con la serie, es decir, el ciclo, $(x_t - x_t^{tr})^2 = (x_t^c)^2$, y la suma de la diferencia al cuadrado entre las tasas de crecimiento de la tendencia consecutivas entre sí, ponderado por un parámetro λ. En otras palabras, para cada periodo la primera sumatoria castiga por desviaciones muy grandes con la serie, la segunda sumtoria castiga por cambios bruscos en la pendiente de la tendencia y λ determina la importancia relativa de tener una tendencia "suave" versus una que se ajuste bien a la serie observada. Notemos que a medida que λ es mayor, la tendencia se vuelve cada vez más suave, con dos casos extremos:

- $\lambda \to \infty$: Tendencia lineal.
$$\Delta x_{t+1}^{tr} = \Delta x_t^{tr}, \quad \forall t$$
- $\lambda = 0$: Tendencia completamente igual a la serie.
$$x_t = x_t^{tr}, \quad \forall t$$



En la práctica, la fijación de λ es arbitraria. Hodrick y Prescott sugierieron en su momento $\lambda = 1600$, pero diversas revisiones de la literatura han hecho críticas a este valor. Con todo, el filtro HP es el más utilizado para estimar el ciclo económico.

2.2. El costo de los ciclos económicos

Lucas (1988) se hizo la pregunta de cuánto aumentaría el bienestar si se lograsen eliminar los ciclos económicos.

2.2.1. Lucas (1988)

Asumiendo que el logaritmo del consumo per cápita (interpretado como el de un hogar representativo) puede ser descompuesto en la suma de una tendencia lineal (con tasa de crecimiento g) con un componente cíclico i.i.d. normal con media $-\frac{1}{2}\sigma_D^2$ y varianza σ_D^2. Es decir,

$$C_t = (1+\lambda)(1+g)^t e^{\varepsilon_t^D} \implies \log C_t = \log(1+\lambda) + t \cdot \log(1+g) + \varepsilon_t^D$$

Donde $E[e^{\varepsilon_t^D}] = 1$.[7] Definiremos una función de valor

$$W(\lambda, \beta, g, \gamma, \sigma_D) = E_0\Big[\sum_{t\geq 0} \beta^t u(C_t)\Big]$$

Donde la función de utilidad es una CES: $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$. Es decir, γ representa el coeficiente de aversión al riesgo. Bajo este modelo, Lucas pronone calcular el costo del ciclo económico mediante determinar el valor de λ que resuelve

$$W(\lambda, \beta, g, \gamma, \sigma_D) = W(0, \beta, g, \gamma, 0) \tag{2.2}$$

Esto es, λ representa la fracción de consumo que debe aumentar (todos los periodos) para compensar la volatilidad del consumo, σ_D. La parte derecha de la ecuación (2.2) representa el bienestar descontado esperado cuando el consumo no tiene volatilidad ($\sigma_D = 0$). El lado izquierdo representa el bienestar descontado esperado cuando el consumo tiene volatilidad, pero es compensado por una fracción λ más de consumo. Entonces,

$$\text{Con volatilidad: } C_t = (1+\lambda)(1+g)^t e^{\varepsilon_t^D}$$
$$\text{Sin volatilidad: } C_t = (1+g)^t$$

Se puede comprobar que

$$\lambda = e^{\frac{1}{2}\gamma\sigma_D^2} - 1$$

Con lo cual podemos estimar el λ mediante estimar el σ_D con un modelo OLS. Así, para $\gamma = 1$, Lucas obtiene que $\lambda = 0{,}068\,\%$. Este resultado es problemático, pues nos dice que el costo del ciclo es casi irrelevante. El problema de este modelo es que se asume una tendencia determinística.

2.2.2. Lucas con tendencia estocástica

Ahora supondremos que el consumo tiene una tendencia estocástica. En particular, digamos que sigue un camino aleatorio:

$$C_t = (1+\lambda)(1+g)^t[\varepsilon_t \varepsilon_{t-1} \dots \varepsilon_0]C_0 \implies \log C_t = \log(1+\lambda) + t \cdot \log(1+g) + \log C_0 + \sum_{k=0}^{t} \log(\varepsilon_k)$$

[7] Aquí ocupamos que si $\log X$ es $\mathcal{N}(\mu, \sigma^2)$, entonces $E[X] = e^{\mu + \frac{1}{2}\sigma^2}$.

Tomando la primera diferencia y simplificando términos:

$$\Delta \log C_t = t \cdot \log(1+g) - (t-1) \cdot \log(1+g) + \sum_{k=0}^{t} \log(\varepsilon_k) - \sum_{k=0}^{t-1} \log(\varepsilon_k)$$
$$\Delta \log C_t = \log(1+g) + \log(\varepsilon_t)$$

Ahora podemos obtener que

$$\lambda = \left[\frac{1-\beta(1+g)^{1-\gamma} e^{-\frac{1}{2}\gamma(1-\gamma)\sigma_D^2}}{1-\beta(1+g)^{1-\gamma}} \right] - 1$$

Considerando $\gamma = 1$ y $\beta = 0{,}96$, se estima que $\lambda = 0{,}241\,\%$. Esto es un poco más cercano a lo que esperaríamos en la realidad, sin embargo, sigue siendo muy bajo.

2.3. Expectativas racionales (Muth, 1961)

Normalmente, entendemos como expectativas racionales el hecho de que los agentes conocen el modelo subyacente en cuestión, pero es mucho más que esto. En esta sección trabajaremos un concepto un poco más matizado.

Utilizaremos el modelo de Muth (1961), centrado en los mercados de *commodities*. Diremos que la oferta agrícola en el periodo t, q_t^S, depende del precio esperado por los productores en el periodo anterior:

$$q_t^S = c p_t^e + v_t$$

donde v_t es i.i.d. e independiente de precios anteriores. La demanda, por su parte, está dada por

$$q_t^D = 1 - p_t$$

Para encontrar el equilibrio del modelo, necesitaremos asumir alguna forma para la expectativa de los precios.

2.3.1. Expectativas ingenuas

Una primera aproximación (ingenua) sería plantear que cada periodo los productores esperan el último precio observado:

$$p_t^e = p_{t-1}$$

Resolviendo el equilibrio, $q_t^S = q_t^D$, llegamos a que

$$c p_{t-1} + v_t = 1 - p_t$$
$$p_t = 1 - c p_{t-1} - v_t$$

Entonces, p_t seguiría un AR(1), que sería estacionario solo cuando $|c| < 1$. Con este resultado surgen dos problemas:

- Los productores asumen $p_t^e = p_{t-1}$. Sin embargo, una vez que sucede t, se deberían dar cuenta que su expectativa estaba mal, pero no lo hacen. En otras palabras, las expectativas del precio no son congruentes con el proceso que sigue, no son racionales.
- Nada asegura que $|c| < 1$. Es decir, el precio podría divergir.

2.3.2. Expectativas racionales

Las expectativas sobre el precio serán racionales cuando los productores ocupan toda la información disponible:

$$p_t^e = E_{t-1}[p_t]$$

Entonces, en equilibrio,

$$cE_{t-1}[p_t] + v_t = 1 - p_t$$

Tomando la esperanza en $t-1$ a ambos lados nos queda que

$$\begin{aligned}
E_{t-1}[cE_{t-1}[p_t] + v_t] &= E_{t-1}[1 - p_t] \\
cE_{t-1}[p_t] + E_{t-1}[v_t] &= 1 - E_{t-1}[p_t] \\
(1+c)E_{t-1}[p_t] &= 1 \\
E_{t-1}[p_t] &= \frac{1}{1+c}
\end{aligned}$$

Sustituyendo en el equilibrio nos queda que

$$\begin{aligned}
p_t &= \frac{1}{1+c} - v_t \\
q_t &= \frac{c}{1+c} + v_t
\end{aligned}$$

Notemos que ahora los productores asumirán que $p_t = \frac{1}{1+c} - v_t$, por lo que el valor esperado del precio es congruente con su modelo subyacente ($E_{t-1}[p_t] = \frac{1}{1+c} - 0$). Además, podemos relajar el supuesto i.i.d., asumiendo un proceso AR(1) en cambio:

$$v_t = \varphi v_{t-1} + e_t$$

donde $|\varphi| < 1$ y e_t es i.i.d. En equilibrio,

$$cE_{t-1}[p_t] + v_t = 1 - p_t$$

Tomando la esperanza en $t-1$ a ambos lados nos queda que

$$\begin{aligned}
E_{t-1}[cE_{t-1}[p_t] + v_t] &= E_{t-1}[1 - p_t] \\
cE_{t-1}[p_t] + E_{t-1}[v_t] &= 1 - E_{t-1}[p_t] \\
(1+c)E_{t-1}[p_t] + \varphi v_{t-1} &= 1 \\
E_{t-1}[p_t] &= \frac{1}{1+c} - \frac{\varphi}{1+c} v_{t-1}
\end{aligned}$$

Sustituyendo en el equilibrio nos queda que

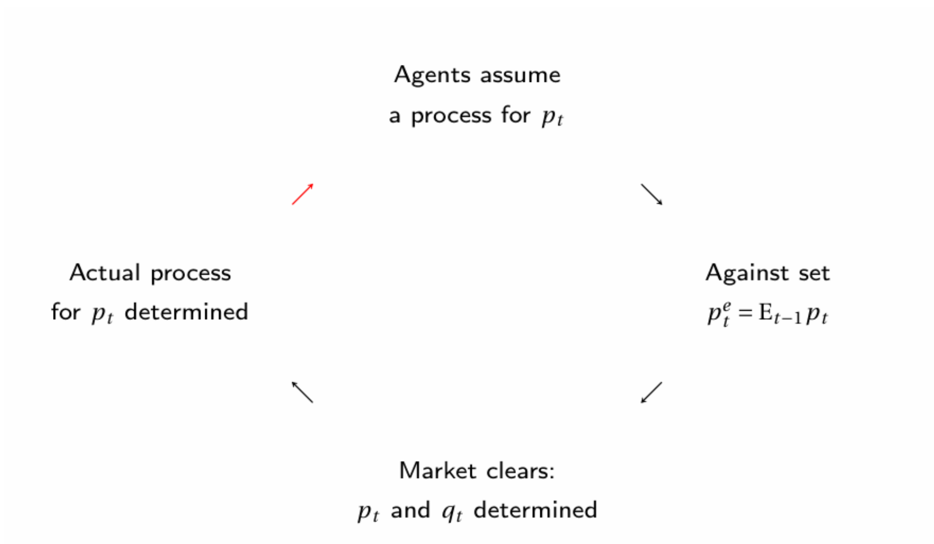
$$\begin{aligned}
\frac{c}{1+c} - \frac{c\varphi}{1+c} v_{t-1} + v_t &= 1 - p_t \\
p_t &= 1 - \frac{c}{1+c} + \frac{c\varphi}{1+c} v_{t-1} - v_t \\
p_t &= \frac{1}{1+c} + \frac{c\varphi}{1+c} v_{t-1} - v_t
\end{aligned}$$

Multiplicando $(1-L\varphi)$ a ambos lados y aplicando que $(1-L\varphi)v_t = e_t$:

$$
\begin{aligned}
(1-L\varphi)p_t &= \frac{1-L\varphi}{1+c} + \frac{c\varphi}{1+c}(1-L\varphi)v_{t-1} - (1-L\varphi)v_t \\
(1-L\varphi)p_t &= \frac{1-\varphi}{1+c} + \frac{c\varphi}{1+c}e_{t-1} - e_t
\end{aligned}
$$

Con lo cual concluimos que p_t sigue un ARMA(1,1). En este caso, de nuevo, los productores asumen que el precio sigue un proceso, en este caso ARMA(1,1), y se forman expectativas acorde a ese proceso, que se confirman a medida que más realizaciones se observan.

2.3.3. Diagrama de expectativas racionales



CONSUMO

3. Problema de Fluctuación del Ingreso

3.1. Formulación general

Tomaremos hogares sujetos a:

- Ingreso exógeno: Y_t.
- Factor de descuento subjetivo: $\gamma = \frac{1}{1+\delta}$.
- Utilidad aditivamente separable: $u(C)$.
- La decisión al inicio de cada periodo de consumir o ahorrar.
- Un activo libre de riesgo con retorno r.

3.1.1. Formulación secuencial

La función objetivo en el periodo t la podemos escribir como

$$\begin{aligned}
\max_{C_t, A_{t+1}} \quad & E_0 \sum_{t=0}^{T-1} \gamma^t u(C_t) \\
s.a. \quad & A_{t+1} = R(A_t + Y_t - C_t), \quad t = 1,..,T \\
& R = 1 + r \\
& A_0 \text{ dado}
\end{aligned}$$

Donde A_t representa la cantidad de activos al inicio del periodo t. Notemos que este problema puede ser entendido como uno de horizonte finito ($T < \infty$), o infinito ($T \to \infty$). Cuando el horizonte es infinito decimos que será necesaria una condición de No-Ponzi. Esta condición de No-Ponzi nos dice que

$$\lim_{T\to\infty} \frac{A_T}{R^T} = 0$$

Esta condición adicional nos permite descartar todas las trayectorias de consumo no realistas (donde el hogar se puede endeudar hasta el infinito, por ejemplo). De manera equivalente, la condición de no-ponzi implica que

$$A_0 = \sum_{t\geq 0} \frac{C_t - Y_t}{R^t}$$

Esta condición asegura que los hogares pagan todas sus deudas.

Demostración. De la restricción presupuestaria intertemporal tenemos que

$$A_0 = R^{-1}A_1 + C_0 - Y_0$$
$$A_0 = R^{-1}(R^{-1}A_2 + C_1 - Y_1) + C_0 - Y_0 = R^{-2}A_2 + \sum_{t=0}^{1} R^{-t}(C_t - Y_t)$$
$$A_0 = R^{-1}(R^{-1}(R^{-1}A_3 + C_2 - Y_2) + C_1 - Y_1) + C_0 - Y_0 = R^{-3}A_3 + \sum_{t=0}^{2} R^{-t}(C_t - Y_t)$$
$$\vdots$$
$$A_0 = R^{-(T+1)}A_{T+1} + \sum_{t=0}^{T} R^{-t}(C_t - Y_t)$$

Luego, si se cumple la condición de no-ponzi ($\lim_{T\to\infty} \frac{A_T}{R^T} = 0$), se cumplirá que

$$A_0 = \sum_{t\geq 0} \frac{C_t - Y_t}{R^t}$$

3.1.2. Formulación recursiva: ecuación de Bellman

Definamos la esperanza del valor presente de la utilidad individual como una función de la riqueza financiera del periodo t:

$$V_t(A_t) = \max_{C_t} \left\{ u(C_t) + \gamma E_t V_{t+1}(A_{t+1}) \right\}$$
$$s.a. \quad A_{t+1} = (1+r)(A_t + Y_t - C_t)$$

Suponiendo solución interior para C_t, la condición de primer orden del problema se obtiene combinando de CPO del lado derecho de la Ecuación de Bellman con el Teorema de la Envolvente. Diferenciando el lado derecho respecto a A_t nos queda que

$$u'(C_t) = \gamma(1+r)E_t V'_{t+1}(A_{t+1}) \tag{3.1}$$

Luego, aplicando el Teorema de la Envolvente respecto a A_t:

$$V'_t(A_t) = \gamma E_t V'_{t+1}(A_{t+1}) \cdot \frac{\partial A_{t+1}}{\partial A_t}$$
$$V'_t(A_t) = \gamma(1+r)E_t V'_{t+1}(A_{t+1})$$

De las dos expresiones anteriores concluimos que $u'(C_t) = V'_t(A_t)$, y por tanto $u'(C_{t+1}) = V'_{t+1}(A_{t+1})$. Reemplazando esto en (3.1) llegamos a la **Ecuación de Euler**:

$$\boxed{u'(C_t) = \gamma(1+r)E_t u'(C_{t+1})} \tag{3.2}$$

3.2. Previsión perfecta

En este modelo asumiremos que no hay incertidumbre en el ingreso. Entonces, recordando que $\gamma = \frac{1}{1+\delta}$, la ecuación de euler será:

$$u'(C_t) = \frac{1+r}{1+\delta} u'(C_{t+1})$$

De esta expresión podemos caracterizar 3 casos:

- $\delta = r \implies C_t$ constante. Se suaviza completamente el consumo.
- $\delta > r \implies C_t$ decreciente (impaciente).
- $\delta < r \implies C_t$ creciente (paciente).

Asumiento una función de utilidad CES de la forma

$$u(c) = \begin{cases} \frac{c^{1-\theta}}{1-\theta}, & \theta > 0, \theta \neq 1 \\ \log c, & \theta = 1 \end{cases}$$

tendremos que

$$\frac{C_{t+1}}{C_t} = \left(\frac{1+r}{1+\delta} \right)^{1/\theta} = (\gamma R)^{\sigma}$$

Donde $R = 1+r$, $\theta = \sigma^{-1}$ es el Coeficiente de aversión relativa al riesgo (CRRA) y σ es la Elasticidad de sustitución instantánea. tomando el logaritmo de lo anterior tendremos que[8]

$$\Delta \log C_{t+1} = \frac{\log \gamma + \log R}{\theta} \cong \frac{r-\delta}{\theta} = \sigma(r-\delta)$$

Es decir, el crecimiento del consumo depende positivamente de la tasa de interés y negativamente de la tasa de impaciencia.

3.3. Equivalencia Cierta

3.3.1. Intuición

Consideremos un hogar con riqueza inicial $\mathscr{W}_0$ en $t = 0$, que se pueden endeudar y ahorrar a una tasa r, sujeto a una condición no-ponzi y motivados por el deseo fundamental de suavizar consumo. Este hogar maximiza el consumo que puede obtener siempre, C. Para consumir C en t, el hogar necesita haber apartado $\frac{C}{(1+r)^t} = CR^{-t}$ de su riqueza inicial en el periodo 0. Es decir, se distribuye la riqueza inicial de tal forma que alcance para consumir lo mismo en todos los periodos. Entonces, el valor más alto que toma C debe satisfacer que

$$\sum_{t\geq 0} R^{-t} C = \mathscr{W}_0 = A_0 + \sum_{t\geq 0} R^{-t} E_0 Y_t$$

Notemos que $\sum_{t\geq 0} R^{-t} = \frac{1}{1-R^{-1}} = \frac{1}{1-\frac{1}{1+r}} = \frac{1}{\frac{r}{1+r}} = \frac{R}{r}$. Entonces,

$$C_0 = \ldots = C_t = C = \frac{r}{R} \mathscr{W}_0$$

[8]Notemos que por aproximación de Taylor sabemos que para cualquier x pequeño $\log(1+x) \cong x$ y $\log(1+x)^{-1} \cong -x$.

Esto es, el hogar coloca su riqueza inicial en el banco, y cada periodo gasta el interés de su ahorro. Luego, este deseo de suavizar consumo conlleva que los hogares reaccionan a las innovaciones en el ingreso de la siguiente manera:

- Si es una innovación de un periodo, sin persistencia (**shock transitorio**): El hogar ahorrará el ingreso extra y aumentará su consumo cada periodo en el interés que genera este ahorro extra.
- Si es una innovación que persiste para siempre (**shock permanente**): El hogar aumentará su consumo cada periodo en una proporción uno-a-uno con el shock.

3.3.2. Análisis formal

Haremos los siguientes supuestos:

1. Utilidad cuadrática:
$$u(c) = c - \frac{b}{2}c^2$$
Es decir, utilidad marginal lineal.
2. $r = \delta \implies \gamma = R^{-1}$.
3. Horizonte infinito.

De la Ecuación de Euler en (3.2) tendremos que

$$u'(C_t) = E_t u'(C_{t+1}) \implies C_t = E_t C_{t+1}$$

Por la Ley de Expectativas Iteradas[9] tendremos que

$$C_t = E_t C_{t+1} = E_t[E_{t+1}C_{t+2}] = E_t C_{t+2} = E_t[E_{t+2}C_{t+3}] = \cdots = E_t C_{t+k}, \quad k = 1, 2, 3, ...$$

Es decir,

$$C_t = E_t C_{t+k}$$

Luego, el error de proyección para $t+1$ será

$$\varepsilon_{t+1} = C_{t+1} - E_t C_{t+1}$$

Pero $C_t = E_t(C_{t+1})$. Entonces podemos plantear el consumo como un camino aleatorio:

$$C_{t+1} = C_t + \varepsilon_{t+1}$$

Este es el conocido como "Camino aleatorio de Hall". A continuación, derivaremos una expresión explícita para C_t y ΔC_t.

[9] Para cualquier set de información I y J, con $I \subset J$, se tendrá que

$$E[E[y|J]|I] = E[y|I]$$

De la condición de no-ponzi sabemos que $A_0 = \sum_{t\geq 0} \frac{C_t - Y_t}{R^t}$ y $R^{-1} = \beta \equiv \gamma$. De esto, podemos plantear la restricción presupuestaria efectiva:

$$\sum_{t\geq 0} \beta^t C_t = A_0 + \sum_{t\geq 0} \beta^t Y_t$$

Tomando la esperanza en 0 llegamos a la restricción presupuestaria esperada:

$$\sum_{t\geq 0} \beta^t E_0 C_t = A_0 + \sum_{t\geq 0} \beta^t E_0 Y_t$$

Pero sabemos por el camino aleatorio del consumo que $E_0 C_t = C_0$. Aplicando esto y desarrollando la suma geométrica del lado izquierdo nos queda que

$$C_0 = \frac{r}{R} \left\{ A_0 + E_0 \sum_{s\geq 0} \beta^s Y_s \right\}$$

Generalizando esto a cualquier periodo:

$$\boxed{C_t = \frac{r}{R} \left\{ A_t + \sum_{s\geq 0} \beta^s E_t Y_{t+s} \right\}} \tag{3.3}$$

Al primer término dentro del paréntesis (A_t) lo llamaremos **riqueza financiera**, al segundo ($\sum_{s\geq 0} \beta^s E_t Y_{t+s}$) lo llamaremos **riqueza humana**. Restando C_{t-1} a la expresión anterior llegamos a que

$$\boxed{\Delta C_t = \frac{r}{R} \sum_{u\geq 0} \beta^u \left\{ E_t Y_{t+u} - E_{t-1} Y_{t+u} \right\}} \tag{3.4}$$

Esto es, el cambio en el consumo de un periodo a otro es proporcional al cambio en el ingreso permanente. Defininendo el ingreso permanente como la anualidad de la riqueza humana:

$$Y_t^P \equiv \frac{r}{R} \sum_{u\geq 0} \beta^u E_t Y_{t+u}$$

Tendremos que

$$\Delta C_t = Y_t^P - E_{t-1} Y_t^P$$

3.3.3. Caso particular: Ingreso Y_t i.i.d.

Asumamos que el ingreso sigue un proceso i.i.d. con media μ:

$$Y_t = \mu + e_t$$

Donde e_t es i.i.d. con media cero. Primero, notemos que

$$E_t Y_{t+s} = \begin{cases} Y_t & , \ s = 0 \\ \mu & , \ s \geq 1 \end{cases}$$

Reemplazando esto en (3.4),

$$\begin{aligned}
\Delta C_t &= \frac{r}{R}\left[Y_t - \mu + \sum_{u\geq 1} \beta^u \{\mu - \mu\}\right] \\
\Delta C_t &= \frac{r}{R}(Y_t - \mu) \\
\Delta C_t &= \frac{r}{R} e_t
\end{aligned}$$

Es decir,

$$\frac{\partial \Delta C_t}{\partial e_t} = \frac{r}{R} < 1$$

Y frente a shocks en el ingreso el consumo se ajusta poco (menos que uno-a-uno).

3.3.4. Caso particular: Ingreso Y_t AR(1)

Ahora asumamos que el ingreso sigue un proceso AR(1) de la forma:

$$Y_t - \mu = \psi(Y_{t-1} - \mu) + e_t$$

Donde $\psi \in [0,1]$. Podemos notar que a medida que aumenta ψ los shocks al ingreso son más persistentes. Además:

- $\psi = 0$ corresponde al caso i.i.d. Todos los shocks transitorios.
- $\psi = 1$ corresponde a un camino aleatorio. Todos los shocks permanentes.

Notemos que, además,

$$E_t[Y_{t+s} - \mu] = \psi^s(Y_t - \mu) \implies E_t Y_{t+s} = \mu + \psi^s(Y_t - \mu)$$

Reemplazando esto en (3.4),

$$\begin{aligned}
\Delta C_t &= \frac{r}{R}\sum_{u\geq 0} \beta^u \left\{\psi^u(Y_t - \mu) - \psi^{u+1}(Y_{t-1} - \mu)\right\} \\
\Delta C_t &= \frac{r}{R}\sum_{u\geq 0} \beta^u \psi^u \{(Y_t - \mu) - \psi(Y_{t-1} - \mu)\} \\
\Delta C_t &= \frac{r}{R}\sum_{u\geq 0} \beta^u \psi^u e_t \\
\Delta C_t &= \frac{r}{R} e_t \sum_{u\geq 0} \beta^u \psi^u \\
\Delta C_t &= \frac{r}{R} \cdot \frac{1}{1 - \beta\psi} e_t
\end{aligned}$$

Pero $\beta = R^{-1} \implies \frac{1}{1-\frac{\psi}{R}} = \frac{1}{\frac{1+r-\psi}{1+r}} = \frac{R}{1+r-\psi} \implies \frac{r}{R} \cdot \frac{1}{1-\beta\psi} = \frac{r}{1+r-\psi}$. Entonces,

$$\Delta C_t = \frac{r}{1 + r - \psi} e_t$$

Luego,

$$\frac{\partial \Delta C_t}{\partial e_t} = \frac{r}{1+r-\psi}$$

Es decir, la respuesta del consumo frente a un shock en el ingreso depende de la persistencia del ingreso, y esta varía de r/R (cuando $\psi = 0$) a 1 (cuando $\psi = 1$).

3.3.5. Dinámica de activos

PENDIENTE

3.4. Teoría del ciclo de vida de Modigliani

En esta teoría, se toma la jubilación como principal motivo para ahorrar. La historia de esta teoría es que la riqueza de los países se va "pasando alrededor":

- Jóvenes tienen poca riqueza. Piden prestado.
- Adultos de mediana edad tienen riqueza. Prestan.
- La riqueza tiene su punto más alto en el momento de jubilación.

A medida que la economía crece más, esta teoría predice que hay más ahorro.

3.4.1. Una formalización simple

Si asumimos $r = \delta = 0$, junto con utilidad instantánea cóncava.

- $r = \delta = 0$.
- Utilidad instantánea cóncava.
- El cohorte nacido en t: tamaño N_t, ingreso per cápita $\bar{y}_t$ en t y 0 en $t+1$.
- Tasas de crecimiento son g para $\bar{y}_t$ y n para N_t.
- Denotamos el consumo en el periodo t como c_t^Y, para la juventud, y c_t^O, para la vejez.

El suavizamiento del consumo, junto a que no existe impaciencia ni interés, nos lleva a que

$$c_t^Y = \frac{1}{2}\bar{y}_t, \quad c_t^O = \frac{1}{2}\bar{y}_{t-1}$$

Denotemos el producto agregado como

$$Y_t = N_t \bar{y}_t \tag{3.5}$$

El consumo agregado como

$$C_t = \frac{1}{2}N_t\bar{y}_t + \frac{1}{2}N_{t-1}\bar{y}_{t-1} = \frac{1}{2}\left\{N_t\bar{y}_t + N_{t-1}\bar{y}_{t-1}\right\}$$

Pero $\bar{y}_{t-1} = \frac{\bar{y}_t}{1+g}$ y $N_{t-1} = \frac{N_t}{1+n}$, entonces

$$C_t = \frac{1}{2}\left\{N_t\bar{y}_t + \frac{1}{(1+g)(1+n)}N_t\bar{y}_t\right\} = \frac{1}{2}\left\{1 + \frac{1}{(1+g)(1+n)}\right\}N_t\bar{y}_t \tag{3.6}$$

Combinando (3.6) y (3.6), llegamos a que el ahorro agregado es

$$S_t = Y_t - C_t = N_t\bar{y}_t - \frac{1}{2}\left\{1 + \frac{1}{(1+g)(1+n)}\right\}N_t\bar{y}_t = \frac{1}{2}\left\{1 - \frac{1}{(1+g)(1+n)}\right\}Y_t$$

Por lo tanto, la tasa de ahorro de la economía, $s \equiv S_t/Y_t$ será

$$s_t = \frac{1}{2}\left\{1 - \frac{1}{(1+g)(1+n)}\right\} \tag{3.7}$$

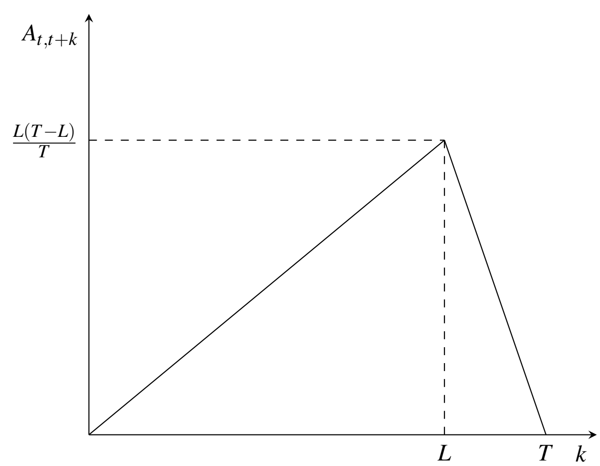
Esto implica que, si bien el ahorro individual de los hogares no depende de g ni n, la tasa de ahorro sí lo hace. Además, según este resultado la tasa de ahorro no depende del tamaño de una economía, sino de su crecimiento. Realizando una expansión de Taylor, tendremos que

$$s \cong \frac{1}{2}(n+g)$$

Si expandiéramos el modelo a que los individuos trabajan L años y viven $T-L$ años adicionales, entonces se puede demostrar que

$$s \cong \frac{1}{2}(n+g)(T-L)$$

Gráficamente,



4. Ahorro por Precaución

En la realidad, cuando se pregunta a los hogares, el principal motivo para ahorrar es para prevenir, por precaución. Sin embargo, modelos como el de equivalencia cierta predicen que ante mayor incertidumbre en los ingresos no cambia el consumo (lo cual es una mala predicción). Entonces, necesitamos un modelo que logre explicar este fenómeno.

4.1. Derivación formal

4.1.1. Dos periodos

Tomemos una primera aproximación donde $T = 2, t = \{0,1\}, r = \delta = 0$ y $A_0 = 0$. No hay motivación de dejar herencia. El hogar resolverá el siguiente problema:

$$\begin{aligned} \max_{C_0, C_1} \quad & u(C_0) + E_0 v(C_1) \\ s.a. \quad & C_0 = Y_0 - A_1 \\ & C_1 = Y_1 + A_1 \end{aligned}$$

Donde Y_1 es estocástico visto desde $t = 0$. Además, permitiremos que las funciones de utilidad difieran entre ambos periodos. Sustituyendo las restricciones en la función objetivo esto nos queda

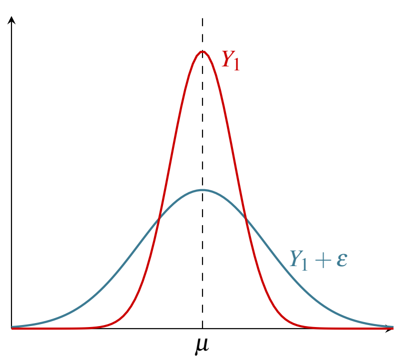
$$\max_{A_1} \quad u(Y_0 - A_1) + E_0 v(Y_1 + A_1)$$

Luego, la condición de primer orden implica la siguiente ecuación de Euler:

$$u'(Y_0 - A_1) = E_0 v'(Y_1 + A_1) \tag{4.1}$$

Notemos que como $u'' < 0$ y $v'' < 0$, el lado izquierdo es creciente en A_1, mientras que el derecho es decreciente en A_1. Asumiendo condiciones de Inada para u y v en cero e infinito[10], podemos asegurar que existe una única solución a esta ecuación.

Ahora consideremos una dispersión manteniendo la media (*mean preserving spread*) de Y_1, que graficamente se vería así:



[10] $u'(0^+) = \infty, v'(0^+) = \infty$ y $u'(\infty) = 0, v'(\infty) = 0$

Denotemos la solución original del problema como A_1^* y la nueva solución de esta dispersión como A_1^{**}. Entonces, habrán tres posibilidades[11]:

1. v' lineal: $A_1^{**} = A_1^{**}$. No cambia el ahorro.
2. v' cóncava: $A_1^{**} < A_1^{**}$. Disminuye el ahorro.
3. v' convexa: $A_1^{**} > A_1^{**}$. Aumenta el ahorro. **Hay ahorro por precaución**. Este es el caso que nos interesa, cuando $v''' > 0$.

4.1.2. Caso general

Ahora exploremos que para cuando

- $T = \infty$.
- $R_t \equiv R$.
- Y_t i.i.d.
- No hay motivación de dejar herencia.
- no existe activo riesgoso.

El hogar resolverá la siguiente ecuación de Bellman:

$$V(A_t, Y_t) = \max_{C_t} \{u(C_t) + \gamma E_t V(R(A_t + Y_t - C_t), Y_{t+1})\}$$

Gracias a que los shocks son i.i.d. podemos traspasar este problema de dos variables de estado a una sola variable de estado: el **efectivo disponible** (*cash-on-hand*).

$$X_t \equiv A_t + Y_t$$

Entonces, podemos reescribir la ecuación de bellman como[12]

$$v(x) = \max_c \left\{u(c) + \gamma E v(x')\right\}$$

Pero $x' \equiv X_{t+1} = A_{t+1} + Y_{t+1} = R(X_t - C_t) + Y_{t+1} \equiv R(x-c) + y'$. Entonces,

$$v(x) = \max_c \left\{u(c) + \gamma E v(R(x-c) + y')\right\}$$

[11]Para llegar a estas posibilidades ocupamos el resultado de expectativas de *mean preserving spreads*. Dado X, ε con $E(\varepsilon) = 0, Var(\varepsilon) > 0$, entonces:

- Si $g(x)$ es estrictamente cóncava: $E[g(X+\varepsilon)] < E[g(X)]$.
- Si $g(x)$ es estrictamente convexa: $E[g(X+\varepsilon)] > E[g(X)]$.

[12]Acá utilizamos el hecho de que el argumento con los valores del periodo t no es relevante, pues se toma la esperanza en t, entonces pasan a ser constantes.

Además, es más conveniente maximizar respecto a los activos del siguiente periodo, $a' = R(x-c)$. Entonces nuestra ecuación de Bellman queda:

$$v(x) = \max_{a'} \left\{ u(x - \frac{a'}{R}) + \gamma E v(a' + y') \right\} \tag{4.2}$$

Luego, tomando la condición de primer orden del lado derecho, llegamos a una ecuación de Euler análoga a la de (4.1):

$$u' \left(x - \frac{a'}{R} \right) = \gamma R E v'(a' + y') \tag{4.3}$$

Entonces, igual que en el caso de dos periodos, podemos imaginar que una dispersión que mantiene la media afecta a y', que va a disminuir el consumo (y aumentar el ahorro por precaución) solo si $v''' > 0$, donde ahora v denota la función valor del problema. asegurar condiciones para que esto ocurra es complicado, sin embargo Sibley (1975) mostró que $u''' > 0$ implica $v''' > 0$. Entonces, nos importará asegurar condiciones bajo las cuales la tercera derivada de la función de utilidad sea positiva. Algunos comentarios:

- A diferencia de u' y u'', ninguna razón teórica nos guía en la elección del signo de u'''.
- Es difícil estudiar u''' a través de los datos.
- En general, el mejor argumento para aceptar $u''' > 0$ es que esto es una condición necesaria para que se cumpla la noción ampliamente aceptada de que la aversión al riesgo es decreciente.
- Las funciones de utilidad CES tienen $u''' > 0$.

4.2. Límite natural de préstamo

Como vimos, el problema de fluctuación del ingreso asume que en todas las trayectoria de ingreso posibles se paga devuelta la deuda. Esto sería diferente si incorporásemos la posibilidad de hacer default en el set de decisiones de los hogares, por ahora no lo haremos. Derivaremos una importante implicancia de que no se pueda hacer default. Primero, este supuesto implica que

$$A_t \geq \sum_{j \geq 0} R^{-j} (C_{t+j} - Y_{t+j})$$

Asumimos que la condición de Inada, $u'(0^+) = \infty$, se cumple. Entonces, los hogares evitarán a cualquier costo los escenarios en los cuales el consumo futuro pueda ser cero. Es decir, $C_{t+j} \geq 0$. Esto implica que

$$A_t \geq -\sum_{j \geq 0} R^{-j} Y_{t+j}$$

Denotando el valor mínimo que puede tomar el ingreso como Y_{min}, podemos asegurar que

$$A_t \geq -\frac{R}{r} Y_{min}$$

En otras palabras, los activos de los hogares tienen que ser mayores o iguales al valor presente descontado del ingreso en el peor de los escenarios posibles. El negativo de esta expresión representa el **límite natural de préstamo** (NBL, por sus siglas en inglés):

$$NBL \equiv \frac{R}{r} Y_{min}$$

Por ejemplo, cuando $Y_{min} = 0$, entonces el límite de endeudamiento de los hogares será $NBL = 0$. Es decir, un hogar nunca se endeuda, porque estaría entrando en una trayectoria donde existe una posibilidad de que el consumo futuro sea cero. Entonces, (4.2) se convierte en

$$v(x) = \max_{a' \geq -NBL} \left\{ u(x - \frac{a'}{R}) + \gamma E v(a' + y') \right\}$$

4.3. Utilidad CARA

La utilidad CARA (*constant absolute risk aversion*) será de la forma

$$u(C) = -\frac{e^{-\alpha C}}{\alpha}$$

Donde se cumple que $u''' > 0$, de modo que habrá ahorro por precaución. Sin embargo, la aversión relativa crece con el ingreso, lo cual no es respaldado por la evidencia micro. Asumiendo $r = \delta = 0$, y un ingreso que sigue un camino aleatorio:

$$Y_t = Y_{t-1} + e_t$$

donde e_t es i.i.d. normal con media cero y varianza σ^2. Los individuos viven T periodos indexados por $t = 0, 1, ..., T-1$. Aplicando la ecuación de Euler de (3.2), llegamos a que

$$e^{-\alpha C_t} = E_t e^{-\alpha C_{t+1}}$$

Se puede mostrar que esta condición de optimalidad es congruente con un camino aleatorio del consumo con *drift* $\frac{\alpha\sigma^2}{2}$:

$$C_t = C_{t-1} + \frac{\alpha\sigma^2}{2} + e_t$$

Además, tenemos la restricción presupuestaria:

$$\sum_{t=0}^{T-1} C_t = A_0 + \sum_{t=0}^{T-1} Y_t$$

Tomando E_0 a ambos lados y usando los resultados del camino aleatorio de Y_t y C_t[13]:

$$E_0 \sum_{t=0}^{T-1} C_t = E_0 \left[A_0 + \sum_{t=0}^{T-1} Y_t \right]$$

...PENDIENTE...

$$C_0 = \frac{A_0}{T} + Y_0 - \frac{\alpha(T-1)\sigma^2}{4}$$

Sustituyendo $T-t$ donde aparece T, obtenemos una expresión para C_t:

$$C_t = \underbrace{\frac{A_t}{T-t} + Y_t}_{C_t^{EC}} - \underbrace{\frac{\alpha(T-t-1)\sigma^2}{4}}_{\text{Ahorro por precaución}}$$

Entonces, derivamos un consumo que es igual al de equivalencia cierta, menos un componente de ahorro por precaución. Este ahorro por precaución:

[13] $E_0 Y_t = Y_0$, $E_0 C_0 = C_0$ y $E_0 C_t = C_0 + t\frac{\alpha\sigma^2}{2}$ para $t = 1, 2, ..., T-1$. Recordar también $\sum_{t=1}^{T-1} t = \frac{(T-1)T}{2}$

- Aumenta con el coeficiente de aversión absoluta al riesgo (α). El hogar se vuelve más cauteloso.
- Aumenta con la varianza de las innovaciones (σ^2). El hogar enfrenta una mayor incertidumbre.
- Aumenta con el tiempo restante de vida (T). El hogar se necesita cubrir más tiempo del riesgo.

Otro posible caso sería cuando Y_t es i.i.d. normal con media $\bar{Y}$ y varianza σ^2. Se puede mostrar que

$$C_t = \frac{r}{R}\left\{A_t + Y_t + \frac{1}{r}\bar{Y}\right\} - P$$

Donde P es una constante que depende de $(r, \alpha, \delta, \sigma)$ y representa el ahorro por precaución. Además, podemos demostrar que cuando $r = \delta$, $P > 0$. Sin embargo, esto implica que los hogares ahorran positivamente todos los periodos (en promedio), y $A_t \to \infty$ con probabilidad uno, lo cual es un problema. Por otro lado, cuando $\delta > r$, generalmente se tendrá una trayectoria de activos acotada.

DERIVAR ESTO DE LA GUÍA 2.

4.4. Buffer-Stock Saving

Este modelo fue desarrollado por Carroll (1997) como una manera de formular el ahorro por precaución. En este modelo, se define el nivel óptimo de efectivo en mano, x^*, como una fracción del ingreso permanente, de manera que

$$E_t[x_{t+1}|x_t = x^*] = x^* \quad \Longleftrightarrow \quad E_t[\Delta x_{t+1}|x_t = x^*] = 0$$

Es decir, el agente buscará mantenerse en una fracción de *cash-on-hand* óptima, x^*. Entonces, cuando $x_t < x^*$:

- El motivo de precaución pesa más que la impaciencia.
- Se acumulan activos (en promedio):

$$E_t[x_{t+1}|x_t < x^*] > x_t \quad \Longleftrightarrow \quad E[\Delta x_{t+1}|x_t < x^*] > 0$$

Por otro lado, cuando $x_t > x^*$:

- La impaciencia pesa más que el motivo de precaución.
- Se desacumulan activos (en promedio):

$$E_t[x_{t+1}|x_t > x^*] < x_t \quad \Longleftrightarrow \quad E[\Delta x_{t+1}|x_t > x^*] < 0$$

En particular, Carroll desarrolla que para algún $\alpha \in (0,1)$:

$$x_{t+1} - x^* \simeq \alpha(x_t - x^*) + e_t \quad \Longleftrightarrow \quad x_{t+1} \simeq (1-\alpha)x^* + \alpha x_t + e_t$$

Donde vemos que la brecha esperada en t entre el efectivo en mano de $t+1$ con el óptimo será una fracción de la brecha en t:

$$E_t(x_{t+1} - x^*) \simeq \alpha(x_t - x^*)$$

Es decir, en promedio, la distancia entre el x efectivo y el x óptimo se reduce a una tasa α cada periodo. Este tipo de modelos son conocidos como **modelos de ajuste parcial**. Este modelo tiene un buen ajuste para los datos de personas que están lejos de su jubilación.

4.5. Restricciones de liquidez

Como sabemos, las restricciones de liquidez consisten en imponer un límite máximo a la deuda. Denotando la deuda en t como D_t, la restricción de liquidez se puede expresar como

$$D_t \leq B$$

Con $B \geq 0$. Notemos que el caso particular en que un hogar no se puede endeudar sería cuando $B = 0$. Como lo habitual es plantearlo en términos de activos ($A_t = -D_t$), normalmente diremos que una restricción de liquidez implica que

$$A_t \geq -B$$

4.5.1. Caso $NBL \leq B$

Supongamos que se cumple la condición de Inada en cero ($u'(0^+) = \infty$). Entonces, si el límite de deuda es mayor o igual al límite natural de préstamo,

$$NBL \leq B$$

La trayectoria óptima del hogar siempre cumplirá la restricción de liquidez. Esto ocurre porque voluntariamente el hogar elige cumplir $A_t \geq -NBL$, y como $NBL \leq B$, podemos concluir que

$$A_t \geq -B$$

Luego, el modelo de *buffer-stock saving* de Carroll puede ser interpretado como un modelo de restricciones de liquidez de $B = 0$. En ese caso, $Y_{min} = 0$, de modo que $NBL = 0$. Este modelo es equivalente a los modelos de ahorro por precaución vistos en esta sección.

4.5.2. Caso $NBL > B$

En este caso, la restrcición de liquidez está activa, pues el hogar eventualmente podría buscar endeudarse más que lo que le permite su restricción de liquidez. Esto ocurre particularmente en los casos en los cuales los hogares no tienen acceso al mercado de crédito ($B = 0$), pero tienen un límite natural positivo ($NBL > 0$).

Por último, si la restricción de liquidez es activa en t, de forma que el consumo en t es menor al que escogería el hogar voluntariamente, la condición de Euler será

$$u'(C_t) > \gamma R E_t u'(C_{t+1})$$

5. Seguros

Hasta ahora, los consumidores solo pueden auto-asegurarse (ahorrar y pedir prestado), no existe diversificación del riesgo.

Una reciente literatura pone el enfásis en la importancia de la diversificación del riesgo para el consumo y el ahorro. Revisaremos el modelo de Townsend (1994) y otros, para revisar esta idea.

5.1. Distribución perfecta del riesgo

5.1.1. Contexto del modelo

Tomaremos un modelo donde hay I hogares ($i = 1, ..., I$), donde la unidad de análisis es una cierta villa de India[14]. En el periodo t, la variable s_t resume el estado de la economía, que determina el producto de todos los hogares.

La estrategia de Townsend (1994) es caracterizar todas las asignaciones Pareto-óptimas y mostrar que estas satisfacen que

$$c_t^i = a^i \bar{c}_t + b^i$$

Es decir, que el consumo de cada hogar i en el periodo t es una función lineal del producto total de la economía (una fracción a^i del consumo total, $\bar{c}_t$, más un componente fijo, b^i). Donde por construcción $\sum_i a^i = 1$ y $\sum_i b^i = 0$.

Luego, para testear el modelo, se estima

$$c_t^i = a^i \bar{c}_t + b^i + \alpha y_t^i$$

Donde si existe perfecta distribución del riesgo, deberíamos estimar $\hat{\alpha} = 0$. Esto es, el consumo de cada hogar no depende de su ingreso (pues se distribuye el riesgo de manera perfecta).

Luego, pensando en todos los periodos, el conjunto de asignaciones Pareto-óptimas resuelven:

$$\begin{aligned} \max_{c_t^i(s^t)} \quad & \sum_{i \in I} \lambda^i \sum_{t=0}^{\infty} \sum_{s^t \in S^t} \gamma^t Pr(s^t) u\big(c_t^i(s^t)\big) \\ s.a. \quad & Y(s_t) \equiv \sum_{i \in I} y^i(s_t) = \sum_{i \in I} c_t^i(s^t) \end{aligned}$$

Donde $\lambda^i > 0$ denota el ponderador del hogar i en la función de bienestar[15] y $s^t = (s_0, s_1, ..., s_t)$ resume la historia de los shocks hasta el periodo t. Notemos que la función objetivo del planificador será

$$\sum_{i \in I} \lambda^i \sum_{t=0}^{\infty} \gamma^t E_0 u(c_t^i)$$

Luego, dados nuestros supuestos de los shocks (aleatorios y no correlacionados), el planificador resuelve una secuencia de problemas estáticos, condicional en s^t, de la forma:

$$\begin{aligned} \max_{c_t^i(s^t)} \quad & \sum_{i \in I} \lambda^i u\big(c_t^i(s^t)\big) \\ s.a. \quad & Y(s_t) = \sum_{i \in I} c_t^i(s^t) \end{aligned}$$

[14]Se analizan economías poco desarrolladas, en las cuales no existe tecnología de almacenamiento, de forma que la variable de estado depende de eventos aleatorios.

[15]Notemos que la sociedad podría ponderar más el bienestar de ciertos hogares (políticos, profesionales, religiosos, etc.).

Como solo $Y(s_t)$ entra en el problema, el planificador puede asumir que $c_t^i(s^{t-1}, s^t) = c_t^i(s_t)$, es decir, que el consumo de cada periodo depende solo de los shocks del último periodo. Entonces, la solución a este problema no exhibe suavizamiento del consumo, contrario a la Hipótesis del Ingreso Permanente de Friedman.

5.1.2. Modelo

El Lagrangeano del problema será

$$\mathscr{L} = \sum_i \lambda^i u\big(c_t^i(s^t)\big) + \mu \sum_i c_t^i(s^t)$$

Townsend asume una función de utilidad CARA, es decir,

$$u^i(c) = -\frac{1}{A^i} e^{-A^i c^i}$$

Luego, la CPO del lado derecho queda:

$$\frac{\partial \mathscr{L}}{\partial c_t^i(s^t)} : \quad \lambda^i e^{-A^i c_t^i(s^t)} = \mu$$

Donde A^i denota el coeficiente de aversión absoluta al riesgo del hogar i. Notemos que esto se cumple para cualquier hogar i, ya que μ no depende del hogar. Entonces, obviando el subíndice t y el argumento s^t, podemos desarrollar lo siguiente para el hogar 1 e i:

$$\begin{aligned}
\lambda^1 e^{-A^1 c^1} &= \lambda^i e^{-A^i c^i} \\
\lambda^1/\lambda^i &= e^{A^1 c^1 - A^i c^i} \\
\log(\lambda^1/\lambda^i) &= A^1 c^1 - A^i c^i
\end{aligned}$$

Definiendo $B^{1,i} \equiv \log(\lambda^1/\lambda^i)$,

$$\begin{aligned}
B^{1,i} &= A^1 c^1 - A^i c^i \\
c^i &= \frac{A^1 c^1 - B^{1,i}}{A^i}
\end{aligned}$$

Pero recordemos que $Y = \sum_i c^i$, entonces,

$$\begin{aligned}
Y = \sum_i c^i &= \left(A^1 \sum_i \frac{1}{A^i} \right) c^1 - \sum_i \frac{B^{1,i}}{A^i} \\
Y + \sum_i \frac{B^{1,i}}{A^i} &= \left(A^1 \sum_i \frac{1}{A^i} \right) c^1 \\
c^1 &= \left(A^1 \sum_i \frac{1}{A^i} \right)^{-1} \left(Y + \sum_i \frac{B^{1,i}}{A^i} \right)
\end{aligned}$$

REVISAR PPT ARREGLADO

6. ¿Por qué Agentes Optimizantes?

6.1. Discusión

Revisar PPT.

6.2. Descuento hiperbólico

6.2.1. Intuición

Existe una extensa literatura que ha documentado que las personas tienden a ser más impacientes con recompensas en un horizonte de corto plazo que en un horizonte de largo plazo.

"La mayoría de las personas prefieren un receso de 15 minutos hoy a uno de 30 minutos mañana. Sin embargo, desde la perspectiva de hoy, la mayoría de las personas **no** prefieren un receso de 15 minutos en 100 días a uno de 30 minutos en 101 días."

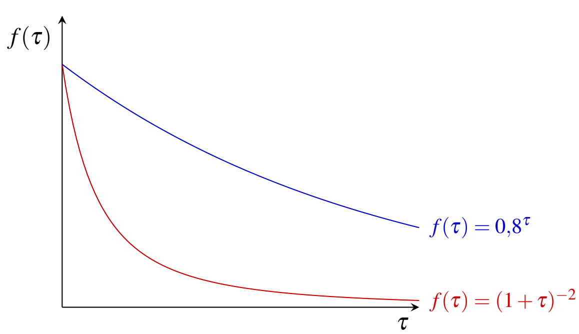
Los agentes podrían darse cuenta de este sesgo, y actuar con anticipación o eliminarse acciones futuras. Entonces, el descuento hiperbólico induce a un problema de auto-control: Quiero ser paciente en el futuro (escoger el receso de 30 minutos), pero en el día 100 voy a preferir una satisfacción instantánea. Matemáticamente, la función de descuento, $f(\tau)$, es definida de manera que si un individuo está indiferente entre consumir c_1 en τ_1 y consumir c_2 en τ_2, entonces,

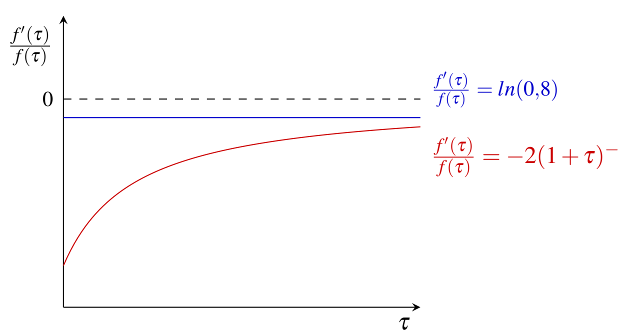
$$c_1 f(\tau_1) = c_2 f(\tau_2)$$

Sin pérdida de generalidad, podemos asumir $f(0) = 1$. Hasta ahora, hemos asumido que la función de descuento es $f(\tau) = \gamma^\tau$, lo cual significa que estar indiferente entre c_1 en τ_1 y c_2 en τ_2 es equivalente a estar indiferente entre c_1 en $\tau_1 + k$ y c_2 en $\tau_2 + k$. Y esto no cumple con lo que documenta la literatura. De hecho, la evidencia sugiere que la función de descuento es aproximadamente una función hiperbólica:

$$f(\tau) = (1 + \alpha\tau)^{-\delta/\alpha}$$

Con $\delta > 0$ y $\alpha > 0$. Notemos que esta función decrece rápidamente para valores pequeños de τ y lentamente para valores grandes. Además, la función de descuento instantánea, $\frac{f'(\tau)}{f(\tau)}$, no será constante como en el caso exponencial. A continuación la intuición gráfica de esta diferencia tomando $\gamma = 0{,}8$, $\delta = 2$ y $\alpha = 1$:





6.2.2. Aplicación económica

Laibson et al. reconoció la dificultad de trabajar con un función de descuento hiperbólica (porque ahora tendríamos otra variable que depende de t), por lo que propuso un factor de descuento “quasi-hiperbólico”:

$$U_t = u(C_t) + \eta \sum_{i=1}^{\infty} \gamma^i u(C_{t+i})$$

Es decir, además, de la tasa de descuento típica, γ, se descuenta todo el consumo futuro por η. Notemos que cuando el descuento es exponencial $\eta = 1$. En cambio, en una economía con descuento hiperbólico $\eta < 1$.

Este descuento hiperbólico también se conoce como “sesgo presente”, porque $\eta < 1$ significa que el consumidor valora el consumo presente de una forma dinámicamente inconsistente. En otras palabras, lo que en el periodo t considera como un consumo óptimo en $t+k$, no será lo mismo cuando lo vuelva a considerar en el periodo $t+k-1$, todo lo demás constante.

En este modelo, existen dos posibilidades de consumidores:

- Descontador hiperbólico **ingenuo**: No tiene en cuenta la inconsistencia dinámica y no hace nada para resolverla.
- Descontador hiperbólico **sofisticado**: Sí tiene en cuenta la inconsistencia dinámica, y escoge su consumo óptimamente hoy, consciente del juego que jugará con sus futuros yos.

6.2.3. Equilibrio recursivo

Pensemos en un individuo que consume 3 periodos, con $\gamma = 1$:

$$U_t = \begin{cases} \log C_1 + \eta(\log C_2 + \log C_3), & \text{si } t = 1 \\ \log C_2 + \eta \log C_3, & \text{si } t = 2 \\ \log C_3, & \text{si } t = 3 \end{cases}$$

Luego, existirá una inconsistencia dinámica entre el periodo 1 y 2, ya que en el periodo 1 el individuo valora igualmente el consumo de 2 y 3, pero en el periodo 2 valora más el consumo de 2.

Si asumimos un consumidor hiperbólico sofisticado con $r = 0$ y $W = 1$, podemos resolver el problema anterior de manera recursiva. Esto es, el “yo” del periodo t escoge su consumo C_t considerando las preferencias de sus futuros yos: Primero, el yo del periodo 3 tiene un rol pasivo, pues solo puede consumir lo que quedó de los periodos pasados:

$$C_3 = 1 - C_1 - C_2$$

Segundo, el yo del periodo 2 toma la riqueza que quedó del periodo 1 y la reparte entre el periodo 2 y 3, acorde a que valora más el consumo inmediato. Es decir, resuelve

$$\begin{aligned}
\max_{C_2} \quad & \log C_2 + \eta \log(1 - C_1 - C_2) \\
\frac{\partial}{\partial C_2}: \quad & \frac{1}{C_2} - \frac{\eta}{1 - C_1 - C_2} = 0 \\
& 1 - C_1 - C_2 = \eta C_2 \\
& C_2 = \frac{1}{1+\eta}(1 - C_1) \\
& \Longrightarrow C_3 = \frac{\eta}{1+\eta}(1 - C_1)
\end{aligned}$$

Estas son las funciones de reacción del yo del periodo 2. Finalmente, el yo del periodo 1 considera el comportamiento de sus yos del periodo 2 y 3 y por lo tanto resuelve lo siguiente:

$$\max_{C_1} \quad \log C_1 + \eta \log \frac{(1 - C_1)}{1+\eta} + \eta \log \frac{\eta(1 - C_1)}{1+\eta}$$

Es decir,

$$\begin{aligned}
\frac{1}{C_1} - \eta \cdot \frac{1+\eta}{1-C_1} \cdot \frac{1}{1+\eta} - \eta \cdot \frac{1+\eta}{\eta(1-C_1)} \cdot \frac{\eta}{1+\eta} &= 0 \\
\frac{1}{C_1} - \frac{\eta}{1-C_1} - \frac{\eta}{1-C_1} &= 0 \\
\frac{1}{C_1} &= \frac{2\eta}{1-C_1} \\
1 - C_1 &= 2\eta C_1 \\
1 &= (2\eta + 1) C_1 \\
C_1 &= \frac{1}{1+2\eta}
\end{aligned}$$

Lo que reemplazando en las funciones de reacción de los yos de los periodos 2 y 3 nos entrega el siguiente conjunto de soluciones:

$$C_1 = \frac{1}{1+2\eta}, \quad C_2 = \frac{2\eta}{(1+2\eta)(1+\eta)}, \quad C_3 = \frac{2\eta^2}{(1+2\eta)(1+\eta)}$$

Este es el equilibrio de un consumidor hiperbólico sofisticado. Notemos que cuando $\eta = 1$, el individuo consume $C_1 = C_2 = C_3 = 1/3$ (porque no tiene impaciencia). En general, a medida que aumenta el η, el individuo consumirá más en los primeros periodos y menos en los últimos.

6.2.4. Equilibrio con compromiso

Otra opción de equilibrio es cuando el consumidor es capaz de comprometerse a niveles futuros de consumo (e.g. poniendo parte de su riqueza en algún instrumento financiero que no le permita retirar antes de tiempo). Con nuestro ejemplo de tres periodos, el individuo podría apartar parte de su riqueza de una forma que no pueda acceder a ella en 2, pero sí pueda en 3. El resultado sería que el individuo planificará en el periodo 1 un consumo igual en el periodo 2 y 3 (porque en el periodo 1 los valora igual), entonces el patrón de consumo quedará:

$$\tilde{C}_1 = \frac{1}{1+2\eta} = C_1$$
$$\tilde{C}_2 = \frac{\eta}{1+\eta} = \frac{1+\eta}{2} C_2 < C_2$$
$$\tilde{C}_3 = \frac{\eta}{1+\eta} = \frac{1+\eta}{2\eta} C_3 > C_3$$

Como podemos ver, en comparación al equilibrio recursivo el equilibrio con compromiso resulta en un menor consumo en el segundo periodo y uno mayor en el tercero. Si pensamos en el ciclo de vida, podemos interpretar esto como que el compromiso lleva a mayores ahorros en la etapa mediana de la vida.

7. Déficits

Los gobiernos se financian a sí mismos a través de una combinación de

- Impuestos.
- Deuda.
- Señoreaje.
- Privatizaciones.
- Utilidades de empresas públicas.

Esto nos lleva a preguntarnos: ¿cómo deberían financiarse los gobiernos?. En esta sección discutiremos la disyuntiva deuda v/s impuestos (de suma alzada).

7.1. Equivalencia ricardiana

7.1.1. El modelo keynesiano tradicional

Bajo el modelo keynesiano tradicional el consumo depende positivamente del ingreso disponible:

$$C = C(YD), \quad YD = Y - T + TR$$

Dado que el gobierno desea financiar un aumento en el gasto (ΔG), si este lo financia a través de impuestos:

$$\Delta^+ G \implies \Delta^+ T \implies \Delta^- YD \implies \Delta^- C$$

Mientras que si lo financia a través de deuda, no aumentan los impuestos, no cambia el ingreso disponible y por tanto tampoco cambia el consumo. Bajo este modelo, es mejor financiar mediante deuda.

Sin embargo, modelos más sofisticados incorporan el hecho de que el consumo depende más de la riqueza que del ingreso disponible. Entonces, el efecto de financiar un aumento de gasto con deuda dependerá de si la deuda del gobierno es incorporada en la riqueza de los privados.

7.1.2. La intuición

Barro (1974) redescubrió la idea planteada originalmente por David Ricardo de que frente a un aumento en la deuda de gobierno los individuos se dan cuenta que en algún punto del futuro deberán pagar más impuestos como forma de financiar esta deuda, y por lo tanto aumentan sus ahorros en anticipación a estos impuestos futuros, dejando el consumo sin cambios.

7.1.3. Derivación informal

Denotemos la deuda de gobierno en términos reales del periodo t, $D(t)$. El gobierno retira esta deuda en algún periodo $t' > t$ y debe pagar $D(t)(1+r)^{t'-t}$.

La riqueza privada vista desde el periodo t aumenta porque los hogares son dueños de los bonos de gobierno, por un monto $D(t)$, pero disminuye en la misma magnitud porque en el periodo t' estos tendrán que pagar los impuestos correspondientes, $T(t')$. Visto desde t:

$$D(t) = \frac{1}{(1+r)^{t'-t}} T(t')$$

Como los dos efectos se cancelan, la riqueza de los individuos no depende del financiamiento de los gastos de gobierno y cambios en el gasto/deuda no afectan el consumo.

7.1.4. Derivación formal

Definiciones:

- Interés r de los bonos de gobierno.
- Deuda $D(t)$, gasto $G(t)$, impuestos de suma alzada $T(t)$, todo en términos reales.
- Los procesos de G e Y son exógenos y sin incertidumbre.
- Tiempo continuo.
- Análisis comienza en $t = 0$.
- Los impuestos son de suma alzada (no distorsionadores) y junto con la deuda son la única fuente de financiamiento del gasto.

El cambio en el déficit de gobierno será igual a la diferencia entre gastos e ingresos:

$$\frac{\partial D}{\partial t}(t) = G(t) - T(t) + rD(t) \tag{7.1}$$

Mientras que el déficit primario es el déficit sin el pago de intereses:

$$G(t) - T(t)$$

Integrando por partes el déficit:

$$\int_0^\infty \frac{\partial D}{\partial t}(t) e^{-rt} dt = D(t)e^{-rt}\Big|_{t=0}^{t=\infty} + r\int_0^\infty D(t)e^{-rt} dt$$

Podemos imponer la condición de no-ponzi ($\lim_{t\to\infty} D(t)e^{-rt} = 0$) y notar que

$$D(t)e^{-rt}\Big|_{t=0}^{t=\infty} = \lim_{t\to\infty} D(t)e^{-rt} - D(0) = -D(0)$$

Reemplazando esto en lo anterior, llegamos a que

$$\int_0^\infty \frac{\partial D}{\partial t}(t) e^{-rt} dt = r\int_0^\infty D(t)e^{-rt} dt - D(0) \tag{7.2}$$

Multiplicando (7.1) por e^{-rt} e integrando de 0 a ∞:

$$\int_0^\infty e^{-rt}\left(T(t)+\frac{\partial D}{\partial t}(t)\right)dt=\int_0^\infty e^{-rt}\left(G(t)+rD(t)\right)dt$$

Reemplazando (7.2) y reordenando llegamos a que

$$D(0)=\int_0^\infty \left[T(t)-G(t)\right]e^{-rt}dt \tag{7.3}$$

Que es la **restricción presupuestaria intertemporal de gobierno**, que nos dice que la deuda de gobierno inicial debe ser igual al valor presente de todos los futuros superávits fiscales.

Por su parte, la **restricción presupuestaria del sector privado** será:

$$\int_0^\infty C(t)e^{-rt}dt=\left[D(0)+F(0)\right]+\int_0^\infty \left[Y(t)-T(t)\right]e^{-rt}dt \tag{7.4}$$

Donde $F(0)$ denota los activos iniciales sin incluir los bonos de gobierno. Si asumimos que el consumidor representativo internaliza (7.3), podemos sustituir esto en (7.4) y obtener la **restricción presupuestaria consolidada del sector privado**:

$$\int_0^\infty C(t)e^{-rt}dt=F(0)+\int_0^\infty \left[Y(t)-G(t)\right]e^{-rt}dt \tag{7.5}$$

Entonces, si tenemos en cuenta que el problema del consumidor es:

$$\begin{aligned}\max_{C(t);t\geq 0}\quad & \int_0^\infty e^{-\delta t}u(C(t))dt\\ s.a.\quad & (7{,}5)\end{aligned}$$

Notamos que, como D y T no aparecen en el problema del consumidor, podemos concluir que la forma en la cual el gobierno financia sus gastos es irrelevante para la definición del consumo óptimo de los privados.

7.1.5. Implicancias

- La deuda pública no es riqueza neta (no aumenta el consumo).
- La decisión entre financiar el gasto de gobierno con impuestos no distorsionadores o deuda es irrelevante para el consumo.
- El tamaño de la deuda pública es irrelevante.
- Una política fiscal contracíclica no tiene ningún efecto.

7.2. Críticas a equivalencia ricardiana

7.2.1. Vidas finitas, entrada de nuevos hogares

Los hogares que compran los bonos de gobierno hoy no son necesariamente los mismos que tendrán que pagar los impuestos futuros para financiar esa deuda. Es decir, **financiar mediante deuda implica hacer a los hogares de hoy más ricos y a los de mañana más pobres**.

Dos respuestas a este argumento:

- Si existe una preocupación de los individuos por sus descendientes, entonces estos actuarán “como si” su horizonte fuera infinito.
- Para efectos prácticos, los individuos viven lo suficiente como para que la equivalencia ricardiana sea una buena aproximación a su comportamiento.

7.2.2. Restricciones de liquidez

Si existen restricciones de liquidez activas, emitir deuda aumenta el consumo, ya que esta funcionará como si el gobierno esté pidiendo prestado por los individuos.

Una respuesta a este argumento:

- Las restricciones de liquidez son endógenas. Por lo tanto, frente a un aumento en la deuda de gobierno, los prestamistas internalizarán un riesgo más elevado y elevarán las restricciones. cancelando el efecto.

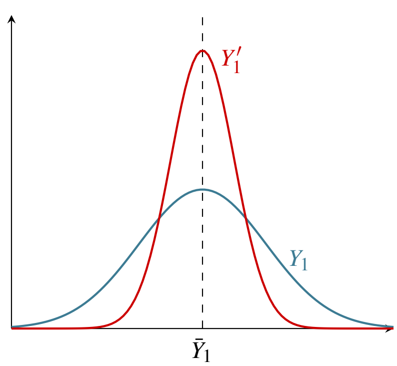
7.2.3. Impuestos distorsionadores y ahorro por precaución

En general, los impuestos crecen con el ingreso. Supongamos ahorro por precaución ($u''' > 0$) y agentes heterogéneos en su ingreso futuro, es decir, el ingreso crece más para algunos que para otros. Digamos que en $t = 0$ todos los individuos tienen un ingreso $\bar{Y}_0$, mientras que en $t = 1$ la distribución de ingresos es una normal con media $\bar{Y}_1$ y varianza σ^2.

Luego, existirán dos escenarios para financiar un aumento en G:

1. Se cobran impuestos en $t = 0$.
2. Se emite deuda en $t = 0$ y se cobran impuestos en $t = 1$ para pagarla.

En el escenario 1 la distribución de ingresos en $t = 1$ es la normal mencionada. Por otro lado, en el escenario 2, al disminuir el ingreso disponible de todos de una forma proporcional, la varianza del ingreso será menor:



Y como vimos en la sección de ahorro por precaución, cuando los hogares enfrentan una menor incertidumbre futura, estos disminuyen su ahorro por precaución y son capaces de aumentar su consumo presente. Entonces, si asumimos que se valora más el consumo presente que el futuro, el escenario donde el gasto se financia con deuda lleva a un mayor bienestar.

7.3. Consumidores *hand-to-mouth*

Si bien la equivalencia ricardiana nos permite modelar de una forma microfundada la relación entre la decisión de consumo de los hogares y la fijación de impuestos, esta forma de entender el consumo choca con la existencia de las reglas fiscales. En general, los países fijan reglas fiscales para ahorrar en tiempos de ingresos altos y gastar en tiempos de ingresos bajos. Sin embargo, el modelo de equivalencia ricardiana implica que estas reglas fiscales no tienen ningún efecto y que la regla óptima es no tener ninguna regla, ya que sea cuál sea la regla los agentes la "deshacen" y esta queda sin efecto.

Entonces, para racionalizar las reglas fiscales necesitamos incorporar algún supuesto. El suspuesto más común es el de los ***hand-to-mouth consumers***, consumidores que se consumen todo su ingreso disponible cada periodo ($C_t = Y_t$).

7.3.1. Campbell y Mankiw (1989)

Este es el modelo precursor de la idea de consumidores *hand-to-mouth consumers*. Asumamos que una fracción λ del ingreso agregado es de individuos que consumen todo su ingreso:

$$Y_{1,t} = \lambda Y_t \implies C_{1,t} = Y_{1,t}$$

Y otra fracción $1-\lambda$ viene de consumidores que siguen un camino aleatorio como el que hemos visto:

$$Y_{2,t} = (1-\lambda)Y_t$$

Entonces,

$$\Delta C_{1,t} = \Delta Y_{1,t} = \lambda \Delta Y_t, \quad \Delta C_{2,t} = e_{2,t}$$

Con $e_{2,t}$ ruido blanco. Luego, el cambio agregado de consumo será:

$$\Delta C_t = \Delta C_{1,t} + \Delta C_{2,t} = \lambda \Delta Y_t + e_{2,t}$$

Que es un modelo que podemos estimar para conocer λ. El problema de esta especificación es que por definición ΔY_t está positivamente correlacionado con $e_{2,t}$. Esto se puede solucionar instrumentalizando los cambios en el consumo pasado ($\Delta C_{t-1}, \Delta C_{t-2}, ...$, ya que por resultado del camino aleatorio sabemos que la innovación presente del consumo no estará correlacionada con los componentes pasados.

Campbell y Mankiw utilizan datos de EE.UU. para estimar lo anterior y encuentran un $\hat{\lambda} \cong 0{,}5$. Es decir, alrededor de la mitad de los consumidores serían *hand-to-mouth*.

7.3.2. Consumidores ricos y *hand-to-mouth*

Kaplan, Violante y Weidner (2014) se dieron cuenta que problablemente no todos los consumidores *hand-to-mouth* eran personas pobres (porque EE.UU. no tiene 50 % de pobreza). Entonces, definieron el concepto de ***wealthy hand-to-mouth households***: Hogares que mantienen poca o nada de su riqueza líquida, a pesar de tener una buena cantidad de activos ilíquidos. Luego, estos hogares se comportan como *hand-to-mouth consumers* en el sentido de que no pueden liquidar sus activos para suavizar consumo y su propensión marignal a consumir es cercana a 1.

Kaplan, Violante y Weidner miran datos de EE.UU. usando un modelo que permite la existencia de este tipo de consumidores y concluyen que:

- Entre $25\,\% - 40\,\%$ de los hogares son *hand-to-mouth*. Dos tercios de estos son ricos y un tercio pobres.
- Los hogares HtM pobres son principalmente hogares jóvenes, mientras que los hogares HtM ricos están en torno a los 40 años.
- Los hogares HtM pobres y ricos exhiben respuestas similares frente a shocks de ingreso, significativamente más grandes que la respuesta de los hogares que no son HtM.

7.4. Suavizamiento de impuestos

7.4.1. Sin incertidumbre

Cuando los impuestos son distorsionadores, entonces la elección entre deuda e impuestos sí es relevante, incluso bajo un modelo de equivalencia ricardiana.

Consideremos un gobierno benevolente que minimiza el valor presente de las distorsiones, que crecen más que proporcionalmente con los impuestos (pensar en el triángulo de Harberger). Definamos $v(T,Y)$ como el valor monetario de la distorsión generada al recolectar T impuestos cuando el ingreso agregado es Y. Además, supongamos que esta distorsión es una función de la carga tributaria:

$$v(T,Y) = Yf(T/Y)$$

Con

$$f(0) = 0, \quad f' > 0, \quad f'' > 0$$

Es decir, las distorsiones son una función creciente y convexa de la carga fiscal. Pensemos en tiempo discreto, con las trayectorias de G e Y exógenas y conocidas. Entonces, el gobierno minimiza el valor presente del costo de las distorsiones sujeto a satisfacer la restricción presupuestaria:

$$\begin{aligned} \min_{T_o,T_1,\ldots} \quad & \sum_{t\geq 0} \beta^t Y_t f(T_t/Y_t) \\ s.a. \quad & \sum_{t\geq 0} \beta^t T_t = D_0 + \sum_{t\geq 0} \beta^t G_t \end{aligned}$$

Donde asumimos que $\delta = r$. Notemos que este es un problema análogo al problema del consumidor estándar, donde $C_t \to T_t$, $Y_t \to G_t$, $A_0 \to D_0$ y en vez de maximizar $u(C_t)$ (una función cóncava), minimizamos $Y_t f(T_t/Y_t)$ (una función convexa). Entonces, podemos aplicar un análogo de la condición de Euler de (3.2) y decir que en el óptimo:

$$\frac{\partial}{\partial T} Y_t f(T_t/Y_t) = c$$

Con c alguna constante. Esto a su vez implica que

$$f'(T_t/Y_t) = c$$

Pero dado que f es convexa y las condiciones de Inada, la pendiente de f solo puede ser constante si:

$$\frac{T_t}{Y_t} = c$$

Es decir, en el óptimo la carga tributaria debe ser constante. Es es el suavizamiento de los impuestos, que entrega una posible explicación a porque los gobiernos se endeudan frente a los diversos shocks que sufren sus países (para minimizar las distorsiones).

7.4.2. Con incertidumbre

Si permitimos que el gasto siga un proceso con incertidumbre y asumimos una función cuadrática para f (un buen supuesto dado cómo crecen las distorsiones a mayor carga tributaria):

$$f(T/Y) = k(T/Y)^2$$

Entonces, análogo al camino aleatorio de Hall, llegaremos a un resultado como el siguiente:

$$\frac{T_t}{Y_t} = E_t\left[\frac{T_{t+1}}{Y_{t+1}}\right]$$

Luego, adaptando el resultado de Hall, concluimos que la carga tributaria seguirá un camino aleatorio. Es decir, no se podrán predecir los cambios en la carga tributaria. Análogo a (**??**), tendremos que

$$\frac{T_t}{Y_t} - \frac{T_{t-1}}{Y_{t-1}} = \frac{r}{R}\sum_{u\geq 0}\beta^u\left\{E_t\left[\frac{G_{t+u}}{Y_{t+u}}\right] - E_{t-1}\left[\frac{G_{t+u}}{Y_{t+u}}\right]\right\}$$

Por lo tanto, si G/Y sigue un camino aleatorio[16], entonces:

$$\begin{aligned}
\frac{T_t}{Y_t} - \frac{T_{t-1}}{Y_{t-1}} &= \frac{r}{R}\sum_{u\geq 0}\beta^u\left\{\frac{G_t}{Y_t} - \frac{G_{t-1}}{Y_{t-1}}\right\}\\
\frac{T_t}{Y_t} - \frac{T_{t-1}}{Y_{t-1}} &= \left\{\frac{G_t}{Y_t} - \frac{G_{t-1}}{Y_{t-1}}\right\}\frac{r}{R}\sum_{u\geq 0}\beta^u\\
\frac{T_t}{Y_t} - \frac{T_{t-1}}{Y_{t-1}} &= \left\{\frac{G_t}{Y_t} - \frac{G_{t-1}}{Y_{t-1}}\right\}\frac{r}{R}\frac{1}{1-R^{-1}}\\
\frac{T_t}{Y_t} - \frac{T_{t-1}}{Y_{t-1}} &= \left\{\frac{G_t}{Y_t} - \frac{G_{t-1}}{Y_{t-1}}\right\}\frac{r}{R}\frac{R}{r}\\
\frac{T_t}{Y_t} - \frac{T_{t-1}}{Y_{t-1}} &= \frac{G_t}{Y_t} - \frac{G_{t-1}}{Y_{t-1}}
\end{aligned}$$

Lo cual implica que si en algún periodo $T = G$, entonces de ese periodo en adelante no existe más déficit. Una forma alternativa de interpretar este resultado es que los superávits y déficits surgen únicamente cuando se espera que G/Y cambie (es decir, cuando no sigue un camino aleatorio). Ejemplos de esto son las guerras y las recesiones, que aumentan G/Y, haciendo que los gobiernos incurran en déficit por cierta cantidad de periodos.

Sin embargo, los resultados de esta teoría no logran explicar los altos déficits de largo en los que incurren muchos países, lo que motiva el considerar elementos de economía política en nuestros modelos.

[16] Recordar que si X_t sigue un camino aleatorio sin drift, entonces $E_t[X_{t+k}] = X_t,\ \forall k = 0, 1, 2, ...$

INVERSIÓN

8. Historia y conceptos básicos de inversión

8.1. Definición

Entenderemos el capital como un **insumo durable que se utiliza en la producción de otros bienes**. Este incluye el capital físico, capital humano, conocimiento, entre otros. La característica más importante de este bien es que es durable.

Por otro lado, la inversión es el proceso mediante el cual el stock de capital se ajusta a su nivel deseado (la compra y venta de bienes de capital).

8.1.1. Relación entre capital e inversión

$$K_t = (1-\delta)K_{t-1} + I_t \quad \Longleftrightarrow \quad I_t = \Delta K_t + \delta K_{t-1}$$

Notación:

- K_t: Capital al final del periodo t.
- I_t: Inversión durante el periodo t.

Secuencia de eventos:

- La firma termina el periodo $t-1$ con capital K_{t-1}.
- Comienza el periodo t.
- El capital se deprecia y pasa a $(1-\delta)K_{t-1}$.
- La firma invierte I_t, su capital pasa a ser:

$$K_t = (1-\delta)K_{t-1} + I_t$$

- Termina el periodo t.

8.2. Mercado de arriendo de capital

Este es el modelo simplificado que utilizamos en pregrado. Denotamos el precio al cual se arrienda el capital como c y el beneficio de la firma en un momento del tiempo como:

$$\pi(K, X_1, ..., X_n) - cK$$

Donde X_i son las variables que la firma toma como dada (por ejemplo, precios). Entonces tendremos que la CPO respecto a K nos lleva a que

$$\pi_k(K, X_1, ..., X_n) = c \tag{8.1}$$

Es decir, la firma arrienda capital hasta que su ingreso marginal es igual al precio de arriendo. Ocupando el Teorema de la función implícita, derivando respecto a c obtenemos que

$$\frac{\partial K}{\partial c} = \frac{1}{\pi_{k,k}(K, X_1, ..., X_n)}$$

Entonces, como el capital tiene retornos decrecientes, $\pi_{k,k} < 0$, y concluimos que la demanda por capital (arrendado) será decreciente en el precio del arriendo.

Los problemas de este modelo son los siguientes:

- No es realista, porque las firmas no suelen arrendar todo su capital, sino que son dueñas de la mayoría.
- No permite modelar fenómenos como la quiebra.
- No es útil para modelos de equilibrio general, pues nadie es dueño del capital.

8.3. Teoría del acelerador

8.3.1. Clark (1917)

Este modelo se construye sobre los siguientes supuestos:

- La tecnología de producción es de proporciones fijas (Leontief) y determina el capital deseado K^*:

$$Y_t = \text{mín}\{L_t, K_t/a\}$$

- El crecimiento del producto es exógeno, no es afectado por las decisiones de inversión.

De la función de producción, tendremos que

$$K_t = aL_t = aY_t$$

Asumiendo $\delta = 0$ y tomando la primera diferencia:

$$\Delta K_t = I_t = a\Delta Y_t$$

Es decir, la inversión es mayor cuando se acelera la tasa a la cual crece el producto.

Este modelo tiene el problema de que se ajusta muy mal a los datos: la inversión es mucho más persistente que el crecimiento del producto.

8.3.2. Acelerador flexible

Este modelo de Clark (1944) y Koyck (1954) presenta la primera noción de fricciones. La inversión ahora toma tiempo:

$$I_t = \sum_{j=0}^{\infty} \gamma_j (K^*_{t-j} - K^*_{t-1-j})$$

Donde K^*_t captura el capital deseado si no existiesen fricciones. Agregando un término de depreciación, estiman

$$I_t = a \sum_{j=0}^{\infty} \gamma_j \Delta Y_{t-j} + \delta K_{t-1} + \varepsilon_t$$

con buenos resultados empíricos.

8.3.3. Problemas con el acelerador

Primero, en la mayoría de los modelos tanto I_t como Y_t son endógenos.

Segundo, no existe ningún rol de los precios (precio de compra del capital, tasa de interés, impuestos, etc.)

Tercero, no considera costos de ajuste.

Cuarto, no considera expectativas sobre la rentabilidad futura del capital.

Quinto, el producto es exógeno, los efectos de la inversión sobre Y se ignoran.

8.4. Teoría General de Keynes

En este momento fue cuando se comento a pensar en el rol de la tasa de interés. Las firmas calculan la tasa interna de retorno que lleva a que el valor presente de los beneficios menos los costos sea igual a cero, e invierte solo si esta tasa es mayor que la tasa de interés corriente.

El rol de las expectativas es importante: la inversión responde menos a la tasa de interés y más a los “espíritus animales” de los emprendedores.

9. Modelo Neoclásico

9.1. Supuestos

Jorgenson (1963), Hall y Jorgenson (1967)

- La firma maximiza el valor de las acciones, entendido como el valor presente de las utilidades, descontadas a la tasa libre de riesgo: Se ignora cualquier problema de agencia.
- Accionistas son neutros al riesgo: Se ignora los efectos del riesgo en el retorno exigido por los inversionistas.
- Las firmas no emiten deuda: Se ignora la decisión de financiamiento deuda v/s capital.
- Comenzamos asumiendo que no se pagan impuestos.
- La firma es tomadora de precios, no hay asimetrías de información.
- Mercados de capitales perfectos. Las firmas emiten las acciones que quieran a la tasa de interés libre de riesgo.

9.2. Condición de optimalidad

Derivaremos la condición de optimalidad a través método de perturbaciones. Supongamos que existe una trayectoria óptima de K_t y que la perturbamos levemente, luego la función objetivo no puede estar mejor que antes, entonces derivamos las condiciones de optimalidad.

Consideremos que nos desviamos de la trayectoria óptima comprando ΔK unidades de capital en t y vendiendo lo que queda de estas unidades en $t+\Delta t$. Entonces, el beneficio adicional que obtiene la firma por estas transacciones será:

$$B = -p_{K,t}\Delta K + \int_t^{t+\Delta t} [\pi(K_s+\Delta K, X_s) - \pi(K_s, X_s)]\, e^{-r(s-t)} ds + \frac{1}{1+r\Delta t}(1-\delta \Delta t) p_{K,t+\Delta t}\Delta K \tag{9.1}$$

Denotemos la integral de (9.1) por $\mathscr{I}$. Un desarrollo de Taylor de primer orden de K nos lleva a que

$$\pi(K_s+\Delta K, X_s) \cong \pi(K_s, X_s) + \pi_K(K_s, X_s)\Delta K$$

Sustituyendo esto en $\mathscr{I}$:

$$\mathscr{I} = \left[\int_t^{t+\Delta t} \pi_K(K_s, X_s) e^{-r(s-t)} ds\right] \Delta K$$

Luego, reemplazamos $\pi_K(K_s, X_s)$ por $\pi_K(K_t, X_t)$, dado que por un desarrollo de Taylor de primer orden en dos variables de $\pi(K, X)$, el error que genera esta aproximación es proporcional a la variación de K_s y X_s para $s \in [t, t+\Delta t]$, y esta variación será todo lo pequeña que queramos si elegimos un Δt lo suficientemente pequeño. Entonces,

$$\begin{aligned}
\mathscr{I} &\simeq \pi_K(K_t,X_t)\Delta K \int_t^{t+\Delta t} e^{-r(s-t)}ds \\
&= \pi_K(K_t,X_t)\Delta K \left[\frac{1}{-r}e^{rt-rs}\right]\Big|_t^{t+\Delta t} \\
&= \pi_K(K_t,X_t)\Delta K \left[\frac{1}{-r}e^{-r\Delta t} - \frac{1}{-r}\right] \\
&= \pi_K(K_t,X_t)\Delta K \frac{1-e^{-r\Delta t}}{r}
\end{aligned}$$

Pero por aproximación de Taylor de primer orden sabemos que $e^x \simeq 1 + \frac{e^0}{1}x = 1 + x$. Es decir,

$$\mathscr{I} \simeq \pi_K(K_t,X_t)\Delta K \frac{1-1+r\Delta t}{r} = \pi_K(K_t,X_t)(\Delta K)(\Delta t)$$

Sustituyendo esto en (9.1) se obtiene que

$$B = \left[-p_{K,t} + \pi_K(K_t,X_t)(\Delta t) + \frac{1}{1+r\Delta t}(1-\delta\Delta t)p_{K,t+\Delta t}\right]\Delta K$$

Si Δt es pequeño, entonces $\frac{1}{1+r\Delta t} \simeq 1 - r\Delta t$, luego

$$B = [-p_{K,t} + \pi_K(K_t,X_t)\Delta t + (1-(\delta+r)\Delta t)p_{K,t+\Delta t}]\Delta K$$

Notemos que esto aplica tanto para $\Delta K > 0$ como para $\Delta K < 0$. Entonces, como este beneficio adicional debe ser menor o igual a cero en ambos casos, podemos concluir que lo que multiplica ΔK debe ser igual a cero:

$$-p_{K,t} + \pi_K(K_t,X_t)\Delta t + (1-(\delta+r)\Delta t)p_{K,t+\Delta t} = 0$$

Reordenando términos,

$$\frac{p_{K,t+\Delta t} - p_{K,t}}{\Delta t} = (\delta+r)p_{K,t+\Delta t} - \pi_K(K_t,X_t)$$

Tomando el límite cuando $\Delta t \to 0$ y definiendo $\dot{p}_{K,t} \equiv \lim_{\Delta t\to 0}\frac{p_{K,t+\Delta t}-p_{K,t}}{\Delta t}$, obtendremos la condición de optimalidad:

$$\boxed{\pi_K(K_t,X_t) = (\delta+r)p_{K,t+\Delta t} - \dot{p}_{K,t}} \tag{9.2}$$

9.3. El costo de usuario del capital

Si comparamos (8.1) con (9.2) podríamos definir entonces que el costo de arriendo del capital debe ser

$$c \equiv (\delta+r)p_{K,t+\Delta t} - \dot{p}_{K,t} \tag{9.3}$$

Con lo que nuevamente concluimos que

$$\pi_K(K_t,X_t) = c_t$$

Diremos que c_t es el costo de usuario del capital, que es igual al costo de arriendo implícito del capital. Una ventaja de este modelo versus el modelo de arriendo es que ahora sabemos cómo podríamos calcular c_t. Además, c_t se puede generalizar para incorporar diversos impuestos. Por último, es directo ver que c_t es creciente en $p_{K,t}$, r y δ, y decreciente en $\dot{p}_{K,t}$, lo cual tiene sentido económico

9.4. Un caso particular

Por el Teorema de la Envolvente tenemos que π_K en la condición de optimalidad de (9.2) se puede interpretar tanto como una derivada parcial (respecto a K_t) como una derivada total (que incorpora cómo varían los valores óptimos de los X_t con K). Consideremos una función de producción Cobb-Douglas:

$$Y_t = K_t^{\alpha} L_t^{1-\alpha}$$

Donde el trabajo L_t es un insumo variable (en función del insumo fijo K_t) y denotamos mediante $p_{Y,t}$ el precio del bien de producción. Tendremos que los beneficios serán:

$$\pi(K_t, w_t, p_{Y,t}) = p_{Y,t} Y_t - w_t L_t^*(K_t) = p_{Y,t} K_t^{\alpha} [L_t^*(K_t)]^{1-\alpha} - w_t L_t^*(K_t)$$

Donde $L_t^*(K_t)$ es el valor óptimo de trabajo dado K. Entonces, la condición de optimalidad de (9.2), que es la derivada parcial respecto a K será

$$\pi_K = \alpha p_{Y,t} K_t^{\alpha-1} [L_t^*(K_t)]^{1-\alpha} = \alpha p_{Y,t} \frac{K_t^{\alpha} [L_t^*(K_t)]^{1-\alpha}}{K_t} = \alpha p_{Y,t} \frac{Y_t}{K_t}$$

Y si hacemos la derivada total del beneficio, llegaremos a lo mismo. Luego, recordando que $c \equiv \pi_K$, llegamos a que

$$K^* = \alpha \frac{p_{Y,t} Y_t}{c_t}$$

Para pasar de este capital óptimo a la inversión volvemos a usar rezagos distribuidos y decimos que

$$I_t = \sum_{j \geq 0} \gamma_j \Delta K_{t-j}^* + \delta K_{t-1}$$

Dividiendo por K_{t-1} a ambos lados obtenemos una expresión para la tasa de inversión:

$$\begin{aligned}
\frac{I_t}{K_{t-1}} &= \sum_{j \geq 0} \gamma_j \frac{\Delta K_{t-j}^*}{K_{t-1}} + \delta \\
\frac{I_t}{K_{t-1}} &\simeq \sum_{j \geq 0} \gamma_j \frac{\Delta K_{t-j}^*}{K_{t-j-1}^*} + \delta \\
\frac{I_t}{K_{t-1}} &\simeq \sum_{j \geq 0} \gamma_j \log\left(\frac{K_{t-j}^*}{K_{t-j-1}^*}\right) + \delta \\
\frac{I_t}{K_{t-1}} &\simeq \sum_{j \geq 0} \gamma_j \left\{\Delta \log(p_{Y,t-j} Y_{t-j}) - \Delta \log(c_{t-j})\right\} + \delta
\end{aligned}$$

Es decir, el ratio entre la inversión y el capital dependerá positivamente del crecimiento del ingreso y la tasa de depreciación, y negativamente del crecimiento del costo de usuario del capital.

9.5. Evidencia

Cuando Jorgenson estimó las ecuaciones anteriores, es decir, del tipo

$$\frac{I_t}{K_{t-1}} = cte + \sum_{j \geq 0} \gamma_j \left\{\Delta \log(p_{Y,t-j} Y_{t-j}) - \Delta \log(c_{t-j})\right\} + \varepsilon_t$$

Obtuvo valores positivos y significativos para los γ_j, lo que interpretó como un éxito de su teoría. Esta fue la primera teoría formal de inversión que donde los precios tenían un rol.

Sin embargo, Eisner y Nadiri (1968) luego estimaron modelos del tipo

$$\frac{I_t}{K_{t-1}} = cte + \sum_{j\geq 0} \eta_j \Delta \log(p_{Y,t+1-j} Y_{t+1-j}) - \sum_{j\geq 0} \mu_j \Delta \log(c_{t-j}) + \varepsilon_t$$

Y solo encontraron significancia en los η_j, es decir, que los valores significativos de Jorgenson se debían principalmente a los términos del acelerador ($\Delta \log Y$), y no a la presencia de costos de usuario.

En resumen, el modelo neoclásico constituyó un importante avance teórico en la modelación de la demanda por capital en un contexto dinámico. Sin embargo, su ajuste empírico es pobre.

9.6. Aplicación

9.6.1. Efecto de la tasa corporativa en demanda de capital

Queremos estudiar el efecto que tiene la tasa de impuesto a las utilidades de las empresas en la demanda por capital. Denotando por $\tau \in [0,1)$ la tasa de impuestos corporativa, es trivial ver que el valor de K_0 que maximiza

$$(1-\tau)\int_0^\infty [F(K_t, L_t) - wL_t - p_{K,t} I_t]\, e^{-rt} dt \tag{9.4}$$

No dependerá de τ. Este resultado no tiene mucho sentido económico, y no es lo que muchas veces se menciona en la discusión pública.

Entonces, para quebrar la neutralidad de la tasa corporativa, Bustos, Engel y Galetovic (2004) introducen dos supuestos que abren la posibilidad de que la demanda de capital dependa de τ:

1. Los pagos de intereses por la deuda se descuentan de la base imponible (se imputan como costos).
2. La depreciación contable permite reducir la base imponible.

Al contar con la primera ventaja tributaria, las empresas buscarán financiarse lo más posible mediante deuda. Los acreedores por su parte, querrán que parte del financiamiento de la empresa sea mediante capital. Digamos que una fracción exógena $b \in [0,1]$ de la inversión se financia vía deuda.

Además, al contar con la segunda ventaja tributaria, por cada unidad invertida en t, las empresas pueden contabilizar δ_s como costo en $t+s$, donde $\int_t^\infty \delta_s ds = 1$. Entonces, el ahorro en impuestos de la empresa gracias a la depreciación acelerada, producto de invertir 1 en t, será proporcional a

$$z \equiv \int_0^\infty e^{-rs} \delta_s ds \leq 1$$

Tendremos que $z \in [0,1]$, con $z = 1$ cuando se tiene depreciación instantánea.[17]

Ahora, suponemos que la empresa maximiza el valor presente de los dividendos:

$$Div_t = (1-\tau)[F(K_t, L_t) - wL_t - rD_t] + \tau\Delta_t - (1-b)p_t I_t$$

[17] La expresión de (9.4) corresponde al caso donde $b = 0$ y $z = 1$.

Donde $\Delta_t \equiv \int_0^t \delta_{t-s} p_s I_s ds$ es la suma de los descuentos por depreciación permitidos en t, y $D_t \equiv b \int_0^t p_s I_s ds$ es la deuda acumulada en t. Luego, podemos derivar que la condición de optimalidad resulta en

$$\boxed{\pi_K = \frac{1-\tau(b+z)}{1-\tau}\left[(\delta+r)p_{K,t+\Delta t} - \dot{p}_{K,t}\right]}$$

Notando que $0 \leq b+z \leq 2$, sigue que

- $b+z<1$: $\uparrow \tau \implies \downarrow K$.
- $b+z=1$: $\uparrow \tau \implies \Delta K = 0$.
- $b+z>1$: $\uparrow \tau \implies \uparrow K$.

De modo que, si los incentivos para invertir (dadas las ventajas tributarias) son muy fuertes, tal que $b+z>1$, un incremento en la tasa de impuesto corporativa llevaría a un aumento en la demanda por capital.

9.6.2. Evidencia

En general, se encuentran valores de $b+z$ cercanos a uno, junto con que cambios en la tasa de impuestos corporativa tiene un efecto casi nulo sobre la demanda de capital por parte de las empresas.

10. Teoría q

Esta fue la primera teoría de inversión propiamente tal, que determinó conjuntamente la producción e inversión óptimas. En esta teoría el rol clave lo juega el q-marginal:

$$q_t = \frac{dV_t/dK_t}{p_{K.t}}$$

Donde V_t es el valor de la firma y $p_{K,t}$ es el precio del capital. Es decir, q_t es el cambio en el valor de la firma frente a un aumento en el capital, por peso que se gasta en este capital. Si no hay costos de ajustar el capital, es claro que la empresa invertirá hasta que $q_t = 1$ en todo momento. Pero sí existen costos de ajuste, entonces,

- $q_t > 1$: Incentivos para aumentar el capital.
- $q_t < 1$: Incentivos para disminuir el capital.

10.1. Modelo

Diremos que existen costos convexos de instalar y desinstalar capital. La firma enfrenta una función de producción con retornos constantes:

$$Y_t = F(K_t, L_t, z_t)$$

Donde z_t denota shocks de productividad. Por simplicidad asumiremos una Cobb-Douglas sin shocks de productividad, es decir,

$$F(K_t, L_t) = K_t^{\alpha} L_t^{1-\alpha}$$

Por otro lado, una demanda isoelástica:

$$Y_t = P_t^{-1/\eta}$$

De modo que la elasticidad será $\varepsilon_{Y,P} = 1/\eta$.[18]

10.1.1. Función de beneficio con empleo optimizado

Denotando por x_t todos los precios dados, la función de beneficios que optimiza sobre L es

$$\pi(K_t, x_t) = \max_{L_t} p_{Y,t} F(K_t, L_t, z_t) - w_t L_t$$

Tomando las formas funcionales mencionadas:

$$\pi(K_t, x_t) = C(x_t) K_t^{\frac{\alpha(1-\eta)}{\eta+\alpha(1-\eta)}}$$

Donde $C(x_t)$ es una constante. Luego, bajo competencia perfecta ($\eta = 0$), $\pi(K_t, x_t) = C(x_t) K_t$ y los retornos son constantes. Por otro lado, con poder de mercado ($\eta > 0$), $\pi(K_t, x_t) = C(x_t) K_t^{\gamma}$ con $\gamma = \frac{\alpha(1-\eta)}{\eta+\alpha(1-\eta)} < 1$ y los retornos son decrecientes.

[18]Notemos que el caso donde $\eta = 0$ corresponde a competencia perfecta, mientras que con $\eta > 0$ existe poder de mercado.

10.1.2. Costos de ajuste convexos

Cuando la firma invierte I paga un costo $p_K I$ más un costo de ajuste $C(I,K)$ en unidades de capital. Es decir, el costo total de ajustar capital será $p_K C(I,K)$. Esta función de costo de ajuste satisface que

- $C(0,K) = 0$.
- $C(I,K) > 0$ para cualquier $I \neq 0$.
- $C_I(i,K) = 0$.
- $C_K(i,K) = 0 < 0$ para todo I.
- $C_{II}(i,K) > 0$ para cualquier $I \neq 0$.

La función más simple que cumple estas condiciones es la de **costos cuadráticos**:

$$C(I,K) = \frac{1}{2} b \frac{I^2}{K}$$

Donde el parámetro b captura la magnitud de los costos de ajuste.

10.1.3. Función objetivo

Maximizan el valor presente descontado de los flujos futuros:

$$\begin{aligned} V(K_0) \quad &= \max_{\{I_t, t \geq 0\}} \int_0^\infty \left[\pi(K_t, x_t) - p_{K,t} I_t - p_{K,t} C(I_t, K_t)\right] e^{-rt} dt \\ s.a. \quad &\{x_t, t \geq 0\}, K_0 \text{ dados} \\ &\dot{K}_t = I_t - \delta K_t \end{aligned}$$

Este es un problema estándar de control óptimo: Control: I_t Estado: K_t Co-estado: λ_t

Donde λ_t es el precio sombra de agregar una unidad de capital, es decir, dV_t/dK_t. Entonces: $\lambda_t = p_{K,t} q_t$. Luego, podemos trabajar con el Hamiltoniano:

$$H(K_t, I_t) = \pi(K_t, x_t) - p_{K,t} \left[I_t + C(I_t, K_t)\right] + \lambda_t (I_t - \delta K_t)$$

Del cual conocemos las condiciones de optimalidad.

1. $H_I = 0$:

$$p_{K,t} + p_{K,t} C_I(I_t, K_t) + \lambda_t = 0 \quad \Longrightarrow \quad \boxed{1 + C_I(I_t, K_t) = \frac{\lambda_t}{p_{K,t}} \equiv q_t} \tag{10.1}$$

   Es decir, I_t está definido implícitamente en función de q_t.

2. $-H_K = \dot{\lambda}_t - r\lambda_t$:

$$-\pi_K(K_t, x_t) + p_{K,t} C_K(I_t, K_t) - \lambda_t \delta = \dot{\lambda}_t - r\lambda_t \quad \Longrightarrow \quad \boxed{(r+\delta)\lambda_t = \dot{\lambda}_t + \pi_K(K_t, x_t) - p_{K,t} C_K(I_t, K_t)} \tag{10.2}$$

   Donde en el caso particular en que $C(I,K) = 0$ recuperamos el resultado de Jorgenson, que implica $q_t = 1$ y $\lambda_t = p_{K,t}$.

3. $\lim_{t \to \infty} \lambda_t e^{-rt} K_t = 0$.

10.2. Estado estacionario y dinámica

Agreguemos supuestos para simplificar el análisis:

$$p_{K,t} = 1, \quad \delta = 0, \quad C(I,K) = \frac{1}{2}b\frac{I^2}{K}, \quad x_t = x$$

Entonces, $\dot{p}_{K,t} = 0$, $\dot{K}_t = I_t$, $\lambda = q$ y π_K solo depende de K. Luego, las condicones de optimalidad de (10.1) y (10.2) equivalen a

$$1 + b\frac{I_t}{K_t} = q \quad \Longrightarrow \quad \dot{K} = \frac{q-1}{b}K \tag{10.3}$$

$$r\lambda_t = \dot{\lambda}_t + \pi_K(K_t) + b\frac{I_t^2}{2K_t^2} \quad \Longrightarrow \quad \dot{q} = rq - \pi_K(K) - \frac{(q-1)^2}{2b} \tag{10.4}$$

Luego, las ecuaciones (10.3) y (10.4) definen la dinámica del modelo. El **estado estacionario**, se dará cuando

$$\dot{K}_t = 0, \quad \dot{q}_t = 0 \tag{10.5}$$

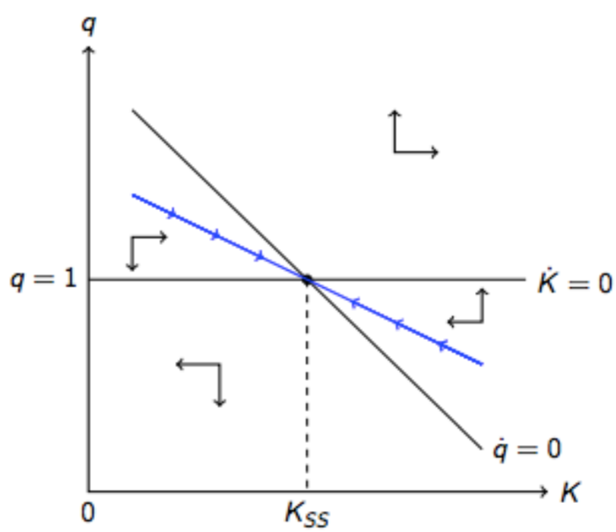
Combinando (10.3), (10.4) y (10.5) junto con el hecho de que $\pi_K = C(x)K_t^{\gamma}$, llegamos a que en estado estacionario

$$q = 1, \quad r = \pi_K = cte$$

Con lo cual el único caso en que el estado estacionario existe es cuando $r = \pi_K$. Para solucionar esto, decimos que el estado estacionario se da cuando

$$q = 1, \quad K = \pi_K^{-1}(r)$$

El supuesto clave detrás de esto es que K se mueve lentamente, mientras que q salta instantáneamente frente a cualquier shock (i.e. es la *jump variable*). Además, la condición de transversalidad se cumplirá solamente si estamos en el brazo estable del sistema, de modo que dado un valor inicial de K_0, el q_0 se ajusta instantáneamente al valor que nos lleva al brazo estable, para luego converger al estado estacionario. Gráficamente:



10.3. Aplicaciones

10.3.1. Shocks de productividad

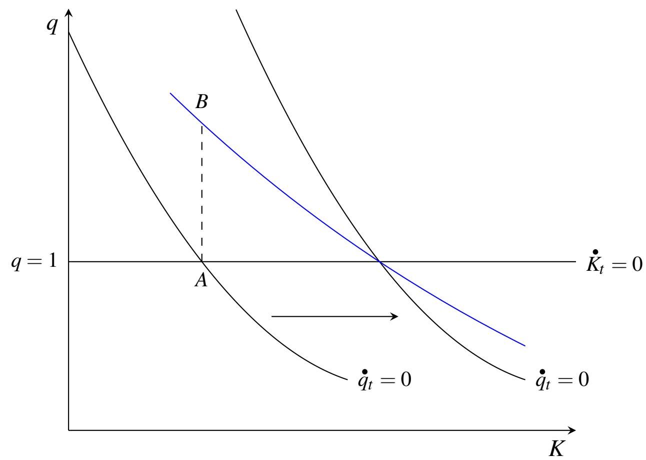
Supongamos que la economía parte en EE, y sucede un shock positivo a la productividad:

$$\pi_K(K) \to z\pi_K(K), \quad z > 1$$

Además, el shock es no anticipado y sucede una sola vez. Entonces, mirando (10.4) tendremos que

$$\dot{q} = rq - z\pi_K(K) - \frac{(q-1)^2}{2b} < 0$$

Y para recuperar el estado estacionario ($\dot{q} = 0$), $\pi_K(K)$ debe caer, es decir, K debe aumentar.



Es decir, inmediatamente después del shock, la economía salta del punto A al B. El valor de q sube, (porque frente al shock de productividad el capital instalado vale más). Luego, el stock de capital comienza a crecer y q va disminuyendo, hasta llegar al nuevo estado estacionario. Es importante que no hayan anticipos al shock.

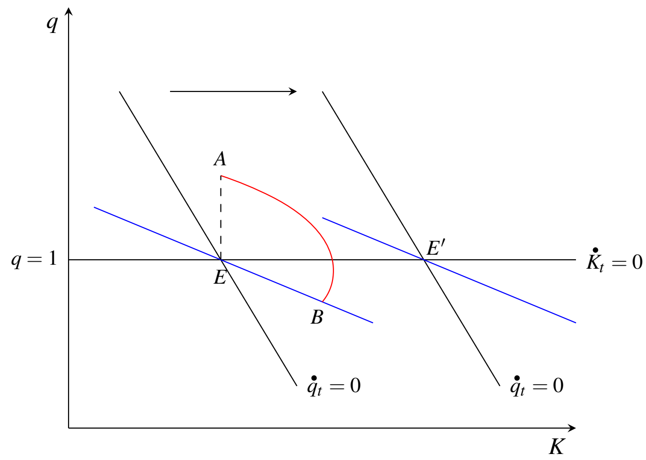
10.3.2. Los MIT shocks

El shock anterior fue “no anticipado y de una sola vez”. Si las firmas lo anticiparan, estas comenzarían a aumentar su capital antes, porque quieren suavizar los costos de ajuste. Si el shock se repitiera, las firmas considerarían la probabilidad de que el shock se repita en cuánto invierten para ahorrar costos de ajuste.

Es decir, para que se cumplan las predicciones del modelo, se requieren supuestos fuertes para protegerse de la crítica de Lucas. Este tipo de shocks teóricos han sido despectivamente catalogados como “MIT shocks”. Sin embargo, a continuación vemos que estos pueden ser útiles para explicar ciertos fenómenos.

10.3.3. Shocks de productividad transitorios

Ahora diremos que en $t = 0$ ocurre el mismo shock que antes, pero ahora este dura hasta $t = T$. Los elementos claves para explicar la dinámica a continuación es que 1. no pueden haber saltos anticipados de q_t y 2. en cada momento manda la dinámica del estado estacionario que está activo:



EXPLICAR GRÁFICO

10.4. Evidencia

10.4.1. q-medio y q-marginal

En los modelos que vimos, el concepto relevante es el q-marginal. Asumiendo $p_{K,t} = 1$:

$$q_t \equiv \frac{dV_t}{dK_t}$$

Sin embargo, esto no es observable. Otra medida que sí se puede medir es el q-medio:

$$\bar{q}_t \equiv \frac{V_t}{K_t}$$

Donde usamos el valor accionario de la firma como V_t.

10.4.2. Resultado de Hayashi (1982)

Supongamos competencia perfecta, de forma que $\pi(K,x) = \pi_K K$, y que los costos de ajuste tienen retornos constantes, de forma que $C(tK,tI) = tC(K,I)$. Entonces:

$$\frac{\partial}{\partial t}: \quad C(K,I) = K \cdot C_K(K,I) + I \cdot C_I(K,I) \tag{10.6}$$

Sin pérdida de generalidad, asumamos $\delta = 0$ y $p_{K,t} = 1$. Luego, aplicando las condiciones que derivamos en (10.3) y (10.4) y que nuevamente $\lambda = q$, tenemos que

$$\begin{aligned} q &= 1 + C_I \\ \dot{q} &= C_K + rq - \pi_K \end{aligned}$$

Luego podemos desarrollar lo siguiente:

$$
\begin{aligned}
(q\dot{K}) &= \dot{q}K + q\dot{K} \\
(q\dot{K}) &= (C_K + rq - \pi_K)K + (1 + C_I)\dot{K} \\
(q\dot{K}) &= (C_K + rq - \pi_K)K + (1 + C_I)I \\
(q\dot{K}) &= K \cdot C_K + I \cdot C_I - K \cdot \pi_K + rqK + I
\end{aligned}
$$

Pero aplicando que tanto C como π tienen retornos constantes (10.6):

$$(q\dot{K}) = C - \pi + rqK + I$$

Multiplicando el lado izquierdo por $e^{-r(s-t)}$ e integrando desde $s = t$ a $s = \infty$:

$$\int_t^{\infty} (q\dot{K})_s e^{-r(s-t)} ds = -q_t K_t + \left[r \int_t^{\infty} (qK)_s e^{-r(s-t)} ds \right]$$

Donde usamos la integración por partes definiendo $dv \equiv (qK)$ y $u \equiv e^{-r(s-t)}$. Haciendo lo mismo al lado derecho:

$$\underbrace{-\int_t^{\infty} \{\pi - C(K_s, I_s) - I_s\} e^{-r(s-t)} ds}_{V_t} + \left[r \int_t^{\infty} (qK)_s e^{-r(s-t)} ds \right] = -V_t + \left[r \int_t^{\infty} (qK)_s e^{-r(s-t)} ds \right]$$

Entonces, combinando ambas igualdades llegamos a que el q-marginal es igual al q-medio:

$$-q_t K_t = -V_t \quad \Longrightarrow \quad \boxed{q_t = \frac{V_t}{K_t} = \bar{q}_t}$$

Una función de costos de ajuste que cumple esta relación es la cuadrática.

10.4.3. Evidencia del resultado de Hayashi

Summers (1981):

- Estima la relación anterior con MCO mediante el siguiente modelo:

  $$\frac{I_t}{K_{t-1}} = \alpha + \beta \bar{q}_t + e_t$$

  Que corresponde a costos de ajuste cuadráticos con $b = 1/\beta$. Recordemos que con esa función de costos de ajuste la condición es

  $$\frac{I_t}{K_t} \frac{1}{b} q - 1$$

- Obtiene un $\hat{\beta}$ muy pequeño (es decir, un $\hat{b}$ muy grande). Específicamente, para un $I/K = 20\,\%$ se tiene que $C(I,K) = 65\,\%$.
- Posibles explicaciones de este mal desempeño del modelo son: 1. No cumplimiento del resultado de Hayashi 2. Errores de medición en $\bar{q}$ ($\hat{\beta}$ sesgado a la baja, $\hat{b}$ sesgado al alza) 3. Endogeneidad y doble causalidad entre I/K y q.

FALTAN AUTORES Y EVIDENCIA

11. Costos no Convexos de Ajuste

La evidencia nos dice que existen 3 tipos de ajuste de capital:

1. Continuo: Mantenimiento.
2. Graduales: refinación y entrenamiento.
3. Grande e infrecuentes.

La teoría neoclásica y q sobre la inversión solo capturan los tipos de ajuste 1 y 2. Necesitamos modelos que capturen el tipo de ajuste de capital grande e infrecuente (especialmente importante para ciertas industrias como la minería).

11.1. Evidencia de ajustes abultados de capital

Doms y Dunne (1998) estudian 12000 plantas manufactureras de EE.UU. y encuentran que los ajustes abultados (lumpy) tienen una relevancia importante. Específicamente, para más de la mitad de las plantas:

$$(1) \quad \max_t I_{it} > 0{,}3 \sum_t I_{it}$$

$$(2) \quad \max_t \Delta K_{it}/K_{it} > 0{,}5$$

Es decir, que un año explica más del 30 % de la inversión histórica (1) y que el año en el que aumentan más su capital lo hacen por más del 50 % (2).

11.1.1. Relevancia macro

Para cada planta i el par (t_i, S_i) está definido por

$$S_i = \max_t I_{it}, \quad t_i =_t I_{it}$$

Que caracteriza el *spike* específico a cada planta. Denotemos por $\mathscr{N}_t$ el número de spikes en el periodo t y por $\mathscr{I}_t$ la inversión promedio de estos spikes. Entonces diremos que la inversión del periodo t se podrá aproximar mediante[19]

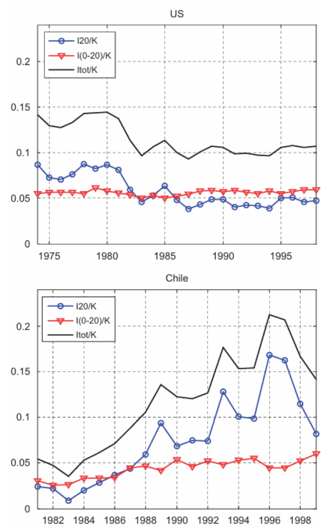
$$I_t \simeq \mathscr{N}_t \mathscr{I}_t$$

Donde $\mathscr{N}_t$ representa el **margen extensivo** (cantidad de empresas que ajustan su capital) y $\mathscr{I}_t$ representa al **margen intensivo** (ajuste promedio de los spikes). La evidencia concluye que el margen intensivo explica mucho mejor la inversión agregada. Es decir, la mayoría de la inversión es explicada por inversiones grandes, más que por cuántas empresas ajustan.

11.2. El modelo más simple con ajuste abultado

El modelo de Calvo. Los agentes ajustan su capital según una probabilidad exógena que está fija entre agentes y en el tiempo.

[19] Detrás de esto está la idea de que las empresas solo ajustan capital mediante spikes, no hay ajustes pequeños o los ajustes pequeños son insignificantes.



11.2.1. Modelo de costos cuadráticos de ajuste

El problema de optimización del agente es

$$\min_{y_{it}} E_t \sum_{j \geq 0} \rho^j \left[c(y_{i,t+j} - y_{i,t+j-1})^2 + (y_{i,t+j} - \hat{y}_{i,t+j})^2 \right]$$

Donde E_t es la esperanza condicional en la información disponible al inicio del periodo t:

$$y_{i,t-1}, y_{i,t-2}, y_{i,t-3}, \ldots; \hat{y}_{i,t}, \hat{y}_{i,t-1}, \hat{y}_{i,t-2}, \hat{y}_{i,t-3}, \ldots$$

$\rho \in (0,1)$ es el factor de descuento y c representa la importancia relativa de la pérdida de ajustarse con respecto a la pérdida de estar fuera del óptimo. Resolviendo, la política óptima viene dada por

$$\boxed{y_{i,t} - y_{i,t-1} = (1-\alpha)\left(y^*_{i,t} - y_{i,t-1}\right)} \tag{11.1}$$

Es decir, cada periodo el agente se ajusta parcialmente $(1-\alpha)$ al target dinámico, que es

$$y^*_{i,t} = (1-\delta) \sum_{k \geq 0} \delta^k E_t \left[\hat{y}_{i,t+k} \right]$$

En otras palabras, el target dinámico ($y^*_{i,t}$) es el valor presente de los target estáticos esperados ($\hat{y}_{i.t}$). Luego, en el agregado:

$$y_t - y_{t-1} = (1-\alpha)(y^*_t - y_{t-1}) \quad \Longrightarrow \quad y_t = \alpha y_{t-1} + (1-\alpha) y^*_t$$

Luego, asumiendo que $\hat{y}$ sigue un camino aleatorio:

$$\hat{y}_t = g + \hat{y}_{t-1} + v_t, \quad v_t \sim i.i.d.(0, \sigma_v^2)$$

Entonces, el target dinámico será

$$y^*_t = (1-\delta) \sum_{k \geq 0} \delta^k (\hat{y}_t + kg) = \hat{y}_t + S$$

donde $S = (1-\delta)g\sum_{k\geq 1} k\delta^k$. Entonces, tenemos el siguiente modelo:

$$y_t = \alpha y_{t-1} + (1-\alpha)(\hat{y}_t + S) = \alpha y_{t-1} + (1-\alpha)\hat{y}_t + (1-\alpha)S$$

Tomando la primera diferencia:

$$\begin{aligned}
\Delta y_t &= \alpha \Delta y_{t-1} + (1-\alpha)\Delta \hat{y}_t \\
\Delta y_t &= \alpha \Delta y_{t-1} + (1-\alpha)(\hat{y}_t - \hat{y}_{t-1}) \\
\Delta y_t &= \alpha \Delta y_{t-1} + (1-\alpha)(g + v_t) \\
\Delta y_t &= (1-\alpha)g + \alpha \Delta y_{t-1} + (1-\alpha)v_t
\end{aligned}$$

Con lo cual concluimos que $\Delta y_t \sim AR(1)$ con el coeficiente de autocorrelación α, una función del costo de ajuste. Es decir, podemos estimar el modelo anterior y recuperar el parámetros de la función de costos de ajuste. En resumen:

- Asumimos una forma para los costos de ajuste (cuadráticos) y un proceso estocástico para el target estático $\hat{y}$.
- Resolvemos el comportamiento óptimo del agente.
- Agregamos para todos los agentes para obtener un proceso ARMA con coeficientes que dependen de los parámetros estructurales.
- Inferimos el parámetro del costo de ajuste estimando el proceso ARMA obtenido.

11.2.2. Modelo de Calvo

Tendremos que la probabilidad de ajustarse en cualquier periodo dado viene dada por π (exógena), independiente entre agentes y a través del tiempo. Si el agente ajusta en el periodo t, escoge $y_{i,t}$ teniendo en cuenta que puede pasar mucho tiempo para que pueda volver a ajustar. Entonces, el agente resuelve:

$$\min_{y_{i,t}} E_t\left[\sum_{k\geq 0}\{\rho(1-\pi)\}^k (y_{i,t} - \hat{y}_{i,t+k})^2\right]$$

Es decir, el agente elige un $y_{i,t}$ tal que se minimice la esperanza del valor presente de la pérdida dada por no cambiar nunca más $y_{i,t}$ (que es el caso relevante). Sigue de este problema que la función a minimizar es cuadrática de la forma

$$Ay_{i,t}^2 - 2By_{i,t} + C$$

donde

$$A = \sum_{k\geq 0}\{\rho(1-\pi)\}^k = [1-\rho(1-\pi)]^{-1}, \quad B = \sum_{k\geq 0}\{\rho(1-\pi)\}^k E_t[\hat{y}_{i,t+k}]$$

Resolviendo se concluye que $y_{i,t}^* = B/A$, es decir,

$$y_{i,t}^* = [1-\rho(1-\pi)]\sum_{k\geq 0}\{\rho(1-\pi)\}^k E_t[\hat{y}_{i,t+k}] \tag{11.2}$$

Para agregar, definimos $y_t = \frac{1}{N}\sum_{i=1}^N y_{i,t}$ y por la Ley de Grandes Números tendremos que una fracción $1-\pi$ de los agentes no ajustará su capital cada periodo, y la otra fracción π sí lo hará. Es decir,

$$y_t = (1-\pi)y_{t-1} + \pi y_t^* \quad \Longrightarrow \quad y_t - y_{t-1} = \pi(y_t^* - y_{t-1})$$

Es decir, en contraste con el modelo de ajuste cuadrático, en el modelo de Calvo el proceso que sigue el agregado es fundamentalmente diferente al que sigue el individual de cualquier agente.

11.2.3. Resultado de equivalencia

Mostraremos que basándonos en los datos agregados, es imposible distinguir entre el modelo de costos de ajuste cuadráticos, donde todos los agentes ajustan continuamente, y el modelo de ajuste abultado, donde solo una fracción de los agentes ajusta en cada periodo. Esta equivalencia fue demostrada por Rotemberg (1987), donde podemos ver que las condiciones derivadas anteriormente para ambos modelos son equivalentes a plantear que

$$\alpha \equiv 1-\pi, \quad \delta \equiv \rho(1-\pi)$$

11.2.4. Irreversibilidad (parcial) del capital

En la realidad π no es exógena, sino que depende positivamente del desequilibrio respecto al capital óptimo. Es decir,

$$\pi = \pi(x), \quad x = y^*_{i,t} - y_{i,t-1}$$

Además, desinstalar capital puede tener un costo mayor a instalar (porque necesitamos encontrar alguien a quien venderlo a un precio razonable). Entonces, el término para $p_K I_t$ que teníamos en la teoría q debe ser reemplazado por un término que es $p_K^+ I_t$ cuando $I_t > 0$ y $p_K^- I_t$ cuando $I_t < 0$, con $p_K^+ \gg p_K^-$. En el caso extremo donde $p_K^- = 0$, tenemos que la **inversión es irreversible**. Cuando $p_K^- > 0$, decimos que existe **irreversibilidad parcial**. La presencia de este problema de irreversibilidad induce a un mayor cuidado por parte de las empresas a la hora de invertir, donde la opción de esperar se vuelve más valiosa mientras más es la brecha entre el precio de compra y el de venta.

11.3. Modelos *Ss* e inversión

11.3.1. Costos no convexos

El ejemplo más simple es pensar en un costo de ajuste fijo dado un nivel de K:

$$C(I,K) = \begin{cases} F, & I \neq 0 \\ 0, & I = 0 \end{cases}$$

Ahora no habrán ajustes pequeños, pues la diferencia entre ajustarse y no ajustarse en fija y grande (F). En muchos periodos será óptimo no ajustarse, y cuando una firma invierte, invierte mucho. Podemos pensar en que el costo fijo depende del nivel de K, por ejemplo, que es proporcional a K. Como la función de costos tiene una discontinuidad en $I = 0$, decimos que son **costos de ajuste no convexos**.

La función de costos anterior puede generalizarse como

$$C(I,K) = F\,\{I \neq 0\} + C_0(I,K)$$

donde $\{I \neq 0\}$ es una *dummy* que toma el valor 1 cuando $I \neq 0$ y 0 si no, y $C_0(I,K)$ es una función de costos de ajuste convexa como las que vimos antes.

11.3.2. Modelo *Ss* sin costos de ajuste

En tiempo discreto, la función de beneficio de la firma será

$$\Pi(K,\theta) = K^\beta \theta - (r+\delta)K$$

donde $0 < \beta < 1$ y θ representa los diferentes shocks que puede enfrentar la firma. Resolviendo para el óptimo estático de capital:

$$K^* = argmax_K \Pi(K,\theta) = \left(\frac{\beta\theta}{r+\delta}\right)^{1/(1-\beta)}$$

Si suponemos que los shocks siguen un camino aleatorio, entonces, $\Delta \log K$ también los sigue.

11.3.3. Introduciendo costos de ajuste

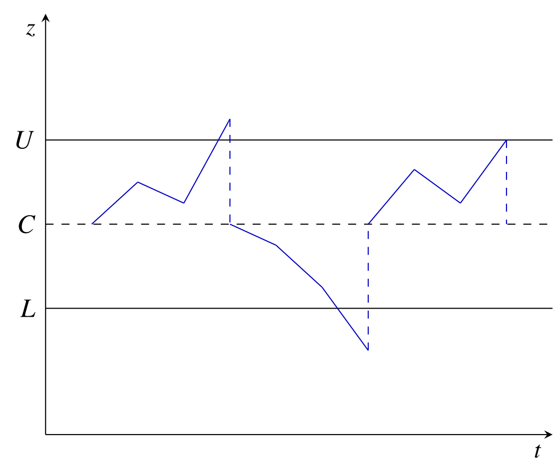
Ahora supongamos que cuando la firma ajusta su capital, deja de percibir una fracción ω de los ingresos de ese periodo. Es decir

$$C(I,K) = \omega K^{\beta}\theta$$

Definamos por z la inversión mandatada por el modelo neoclásico:

$$z = \log(K^*/K) = \log K^* - \log K \equiv k^* - k$$

Luego, resultados generales permiten mostrar que z será la variable de estado y la política óptima será del tipo (L,C,U), donde existe una zona de inacción dentro de la cual el agente no ajusta, y dos límites, superior (U) e inferior (L) que le indican cuándo el agente debe ajustar al óptimo (C).



11.3.4. Limitaciones de las reglas (L,C,U)

Según esta regla, cuando las firmas aumentan su capital, siempre lo hacen en el mismo porcentaje, $U-C$, y cuando lo bajan lo mismo con $C-L$. Esto no es realista, pues en la realidad las diferentes empresas llevarán diferentes trayectorias de z en un mismo periodo. Sin embargo, lo que realmente capturan las reglas (L,C,U) es que la probabilidad de ajustarse es creciente en la variable de desequilibrio de la firma.

DESEMPLEO

12. Salarios de eficiencia

12.1. Mercados Laborales y Macroeconomía

Tres preguntas importantes:

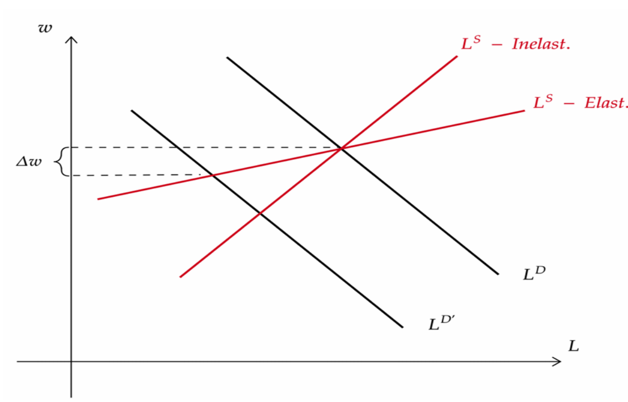
1. ¿Por qué existe desempleo involuntario?
2. ¿Por qué las respuestas del mercado laboral frente a shocks adversos es principalmente a través de cantidades (L) y no vía precios (w)?
3. ¿Por qué las tasas promedio de desempleo varían tanto entre países y a través del tiempo?

12.1.1. ¿Por qué desempleo involuntario?

La respuesta principal son los salarios de eficiencia: Las firmas pagan más que la productividad marginal porque les conviene (asimetrías de información, selección adversa, riesgo moral, etc.). Lo veremos con el modelo Shapiro-Stiglitz.

12.1.2. ¿Por qué cantidades antes que precios?

Si las fluctuaciones del empleo están determinadas por shocks de la demanda por trabajo, podemos decir que la oferta de trabajo es más elástica, lo que contradice la evidencia micro.



12.1.3. Determinantes de la tasa promedio de desempleo

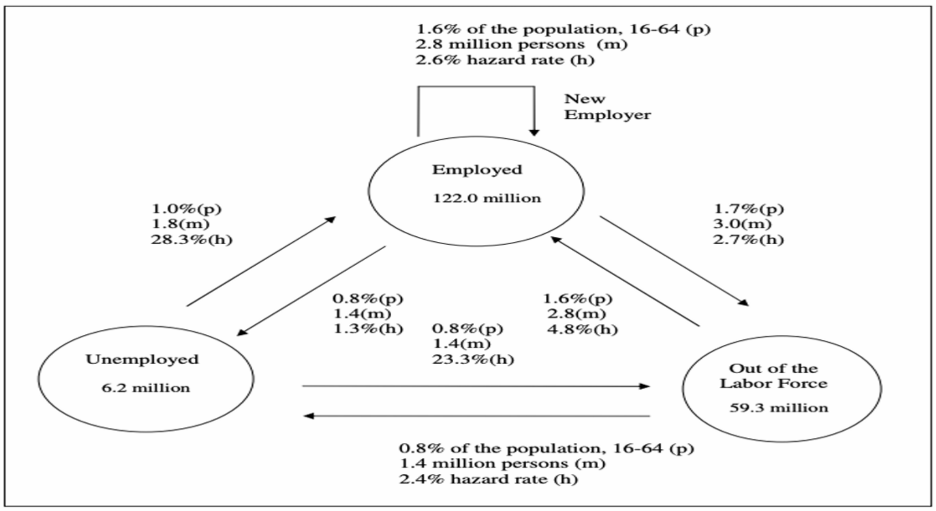
Queremos explicar las diferencias en las tasas promedio de desempleo entre países y dentro de los mismos países a través del tiempo. Para esto, ocuparemos el enfoque de flujos de empleo (creación y destrucción de empleo en cada periodo), donde es necesario modelar las fricciones inherentes al pareo entre firmas y trabajadores que resulta en el empleo. Veremos el modelo de Diamond-Mortensen-Pissiarides (DMP).

12.1.4. Enfoque de flujos en el mercado laboral

Cada periodo hay simultáneamente personas que encuentran empleo y otras que lo pierden. Es decir,

$$\text{Cambio neto de empleo} = \text{Contrataciones} - \text{Separaciones} = \text{Creación} - \text{Destrucción}$$

Figure 1. Average Monthly Worker Flows, Current Population Survey, 1996-2003



12.2. Ecuación de Bellman con shocks exponenciales

Consideremos la decisión de un trabajador de aceptar o rechazar una oferta de trabajo en tiempo continuo. Diremos que el tiempo entre ofertas consecutivas sigue una **distribución exponencial** con parámetro η (tal que el tiempo esperado sea $1/\eta$). Luego, por la ausencia de memoria de la exponencial:

$$Pr(\text{Recibir una oferta entre } t \text{ y } t+\Delta t) = F(\Delta t) = 1 - e^{-\eta \Delta t} \simeq \eta \Delta t$$

Donde en el último paso ocupamos una aproximación de Taylor de primer orden. Es decir, el número esperado de ofertas por unidad de tiempo será η (a un η más grande, más rápido llegan ofertas).

Diremos que el flujo de beneficios en momento s entre ofertas es b_s (subsidio de cesantía). Si se materializa una oferta en t, el agente elige entre aceptar y rechazar. Si acepta, el valor presente descontado esperado de su beneficio es W_t. Si rechaza, sigue recibiendo el flujo b_s hasta la próxima oferta, cuando vuelve a elegir. Entonces la **ecuación de Bellman** del problema es

$$\boxed{rU_t = \dot{U}_t + b_t + \eta \max\{W_t - U_t, 0\}}$$

donde U_t es el valor presente descontado (a tasa r) del beneficio esperado del agente, partiendo en t. La interpretación de esta ecuación es que el agente estará indiferente entre invertir U_t es un bono libre de riesgo con retorno r entre t y $t+\Delta t$, y comprar una unidad del activo en t y venderlo en $t+\Delta t$. Entonces,

$$(r\Delta t)U_t = \dot{U}_t \Delta t + b_t \Delta t + \eta \Delta t \max\{W_t - U_t, 0\}$$

12.3. Modelo Shapiro-Stiglitz

Este modelo de Shapiro y Stiglitz (1984) se basa en el monitoreo imperfecto al que estan sujetas las firmas respecto al esfuerzo de sus trabajadores. Diremos que la utilidad instantánea del trabajador será

$$u(t) = \begin{cases} w(t) - e(t), & \text{si está empleado} \\ 0, & \text{si no} \end{cases}$$

donde $w(t)$ denota el salario y $e(t)$ el esfuerzo del periodo t, que toma solo dos valores, $\bar{e} > 0$ y 0. $\bar{L}$ trabajadores maximizan el valor presente descontado esperado de esta utilidad a una tasa r. N firmas idénticas maximizan el valor presente descontado esperado de su beneficio.

12.3.1. El problema de la firma

Solo los trabajadores que se esfuerzan son productivos, entonces una firma con $L(t)$ trabajadores que se esfuerzan y $S(t)$ que no, tendrá flujo de caja

$$\pi(t) = F(\bar{e}L(t)) - w(t)[L(t) + S(t)]$$

donde $F' > 0$ y $F'' < 0$. Al momento de contratar un trabajador, la firma no sabe si se esforzará o no. Supondremos que

$$\bar{e}F'(\bar{e}\bar{L}/N) > \bar{e} \iff F'(\bar{e}\bar{L}/N) > 1$$

Esto es, que si cada firma contrata una fracción $1/N$ de los trabajadores y todos se esfuerzan, el producto marginal del trabajo excede el costo de esforzarse (el esfuerzo rinde más de lo que cuesta), por lo que con monitoreo perfecto tendremos pleno empleo.

12.3.2. Los trabajadores

Los trabajadores pueden estar en 3 estados:

- E: Empleado y se esfuerza.
- S: Empleado y "flojea".
- U: Desempleado.

Además, puede transitar entre estos estados de la siguiente manera:

- $E \to U$ (Separación fortuita): sigue una variable aleatoria exponencial con parámetro b, donde b es la tasa de separación (exógena).
- $S \to U$ (Separación por monitoreo): sigue una variable aleatoria exponencial con parámetro $b + q$, donde q captura la teconología de monitoreo (a mayor q, mejor es el monitoreo, con $q = \infty$ siendo monitoreo perfecto).
- $U \to E$ (Creación de empleo): sigue una variable aleatoria exponencial con parámetro a, donde a es tomado como exógeno por firmas y trabajadores.

Denotemos por $V_E(t)$, $V_S(t)$ y $V_U(t)$ el valor de la utilidad de un trabajador en cada estado en t. En estado estacionario,

$$\dot{V}_E(t) = 0, \quad \dot{V}_S(t) = 0, \quad \dot{V}_U(t) = 0$$

Entonces, tendremos 3 ecuaciones de Bellman:

$$E: \quad rV_E = (w - \bar{e}) - b(V_E - V_U) \tag{12.1}$$

$$S: \quad rV_S = w - (b+q)(V_S - V_U) \tag{12.2}$$

$$U: \quad rV_U = a(V_E - V_U) \tag{12.3}$$

12.3.3. Interpretación

Considerando la primera ecuación, la interpretación es que un “inversionista” está indiferente entre ser “dueño” de un trabajador empleado que se esfuerza, entre t y $t+\Delta t$, e invertir el mismo monto en un bono libre de riesgo durante ese periodo.

Invertir V_E en un bono libre de riesgo le cuesta V_E y el retorno que obtiene es $r\Delta t V_E$. En cambio, si invierte V_E en un trabajador empleado que se esfuerza, recibirá un dividendo igual a $(w-\bar{e})\Delta t$, pero si el trabajador pierde el empleo (que ocurre con probabilidad $b\Delta t$), tendrá una pérdida de capital igual a $(V_E - V_U)$. Es decir, la pérdida de capital esperada es $b\Delta t(V_E - V_U)$. Igualando estos retornos esperados y desarrollando llegamos a (12.1):

$$\begin{aligned} r\Delta t V_E &= (w-\bar{e})\Delta t - b\Delta t(V_E - V_U) \\ rV_E &= (w-\bar{e}) - b(V_E - V_U) \end{aligned}$$

12.3.4. La condición de No Flojeo

La *no shrinking condition* (NSC) en inglés. La firma pagará el menor salario posible para asegurarse que a los trabajadores les convenga más esforzarse, es decir, $V_E > V_S$. Entonces,

$$V_E = V_S + \varepsilon$$

donde $\varepsilon > 0$ es muy pequeño. Tan pequeño, que podemos suponer $V_E = V_S$ Usando esto al restar (12.2) de (12.1):

$$V_E - V_U = \frac{\bar{e}}{q} \tag{12.4}$$

Notamos que esto va en contra de los modelos Walrasianos, que nos dicen que los trabajadores están indiferentes entre trabajar y no trabajar ($V_E = V_U$). Acá, en cambio, vemos que la diferencia entre la renta de estar empleado (esforzándose) versus desempleado es creciente en el valor productivo del esfuerzo $\bar{e}$ (si me cuesta más esforzarme pediré más compensación) y decreciente en la probabilidad de ser pillado flojeando (mejor monitoreo permite pagar menos).

Restando (12.3) con (12.1) y usando (12.4):

$$\begin{aligned}
-r(V_E - V_U) &= \bar{e} - w + (a+b)(V_E - V_U) \\
w - \bar{e} &= (a+b+r)(V_E - V_U) \\
\frac{w-\bar{e}}{a+b+r} &= V_E - V_U = \frac{\bar{e}}{q} \\
w &= \bar{e} + \frac{a+b+r}{q}\bar{e}
\end{aligned}$$

Es decir, mostramos que

$$w = w(\underset{(+)}{\bar{e}}, \underset{(+)}{a}, \underset{(+)}{b}, \underset{(+)}{r}, \underset{(-)}{q})$$

revisar interpretaciones.

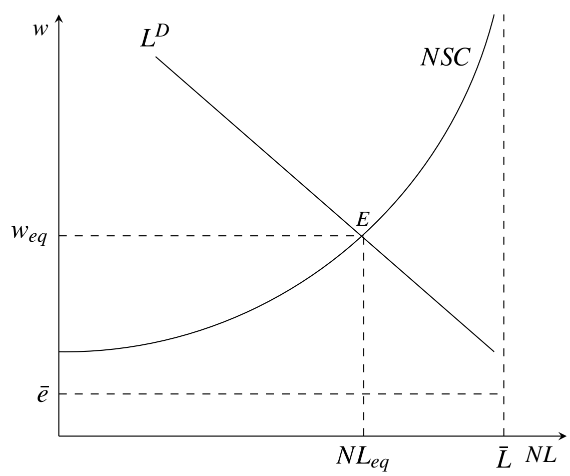
En equilibrio, imponemos que los flujos de empleo a desempleo deben ser iguales a aquellos de desempleo a empleo (Se crean los mismos empleos que los que se destruyen):

$$bNL = a(\bar{L} - NL) \implies a = \frac{bNL}{\bar{L} - NL} \implies a + b = \frac{b\bar{L}}{\bar{L} - NL}$$

Reemplazando en la expresión encontrada para w llegamos a la **condición de no flojeo (NSC)**:

$$\boxed{w = \bar{e} + \left(\frac{b\bar{L}}{\bar{L} - NL} + r\right)\frac{\bar{e}}{q}} \tag{12.5}$$

Que es creciente en NL: Una mayor cantidad de empleos implican que es más fácil encontrar empleo, a los trabajadores no les importa tanto perder el empleo, por lo que el salario que induce esfuerzo crece. Gráficamente:



12.3.5. Análisis del equilibrio

En equilibrio, los trabajadores se esfuerzan, entonces la firma resuelve

$$\max_L F(\bar{e}L) - wL$$

Que nos da la siguiente CPO:

$$\bar{e}F'(\bar{e}L) = w \tag{12.6}$$

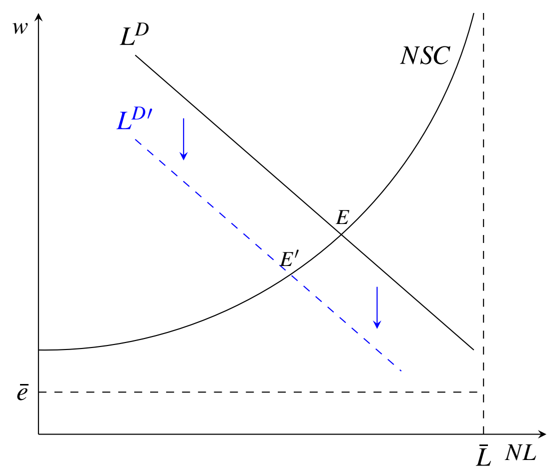
Lo que define una demanda por trabajo correspondiente: $L^D = L^D(w)$. En la intersección de la NSC y L^D está el equilibrio del modelo, donde $w = w_{eq}$ y $L = L_{eq}$.

Notemos que $NL_{eq} < \bar{L}$, porque existen trabajadores que están desempleados involuntariamente (que trabajarían por el salario de equilibrio, pero no encuentran trabajo). En este sentido, el **equilibrio es ineficiente**.

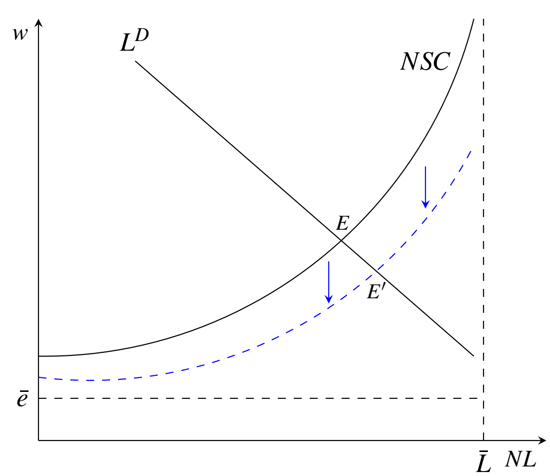
Sin embargo, podemos pensar en el concepto de **equilibrio eficiente dadas las restricciones**, ya que dadas las restricciones del modelo, el equilibrio sigue resultando en una asignación de Pareto.

12.3.6. Shocks de demanda

Si justo el equilibrio del modelo se encuentra en la parte plana de la NSC, un shock a la demanda conllevará un movimiento grande en empleo y pequeño en salarios, que es lo observado en la realidad:



12.3.7. Mejora de la tecnología de monitoreo



13. Modelo Diamond-Mortensen-Pissarides (DMP)

El mercado laboral está sujeto a heterogeneidades en habilidad, fricciones y asimetrías de información. Esto resulta en que los acuerdos de empleo son descentralizados, descoordinados y costosos para ambas partes. A contrario de los modelos walrasianos del mercado laboral, los trabajos generan rentas en equilibrio.

En esta sección, asumiremos que el número de trabajos en cualquier momento del tiempo resulta de una **función de *matching*** que toma en cuenta el número de trabajadores buscando empleo y el número de firmas buscando trabajadores.

13.1. Función de matching

Sea L la fuerza de trabajo, U el número de desempleados y V el número de vacantes, la creación de empleos en un intervalo de tiempo Δt será

$$m(U,V)\Delta t$$

donde m es la función de pareo, que suponemos creciente y cóncava en cada argumento ($m' > 0, m'' < 0$) y con retornos constantes. Luego, definimos 3 nuevas variables: la tasa de vacantes, la tasa de desempleo y la estrechez del mercado laboral:

$$v = \frac{V}{L}, \quad u = \frac{U}{L}, \quad \theta = \frac{v}{u}$$

Cuando θ aumenta, hay más vacantes por desempleados y es más fácil encontrar empleo. Además, hay dos probabilidades instantáneas importantes:

- Tasa a la cual se llenan las vacantes:

$$q(\theta) \equiv \frac{m(uL, vL)}{vL} = m\left(\frac{u}{v}, 1\right) = m\left(\frac{1}{\theta}, 1\right)$$

  Que es decreciente en θ.

- Tasa a la cual los desempleados encuentran empleo:

$$\frac{m(uL, vL)}{uL} = m(1, \theta) = \theta m\left(\frac{1}{\theta}, 1\right) = \theta q(\theta)$$

  Que es creciente en θ.

13.2. Curva de Beveridge

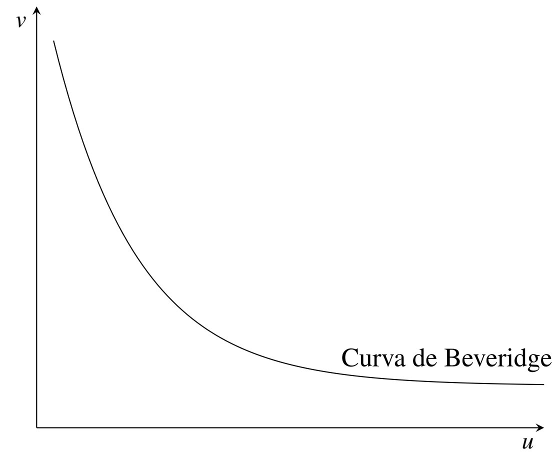
Pensemos en que los shocks exógenos y específicos a cada empleo ocupado llegan a una tasa poisson de λ (trabajadores quedan desempleados, firmas vuelven a sus vacantes, etc.). Entonces, la evolución de la tasa de desempleo será

$$\frac{du}{dt} = \dot{u}_t = \lambda(1 - u_t) - \theta_t q(\theta_t) u_t$$

Donde el primer término representa el número de separaciones y el segundo el número de *matches*, ambos divididos por L. En estado estacionario ($\dot{u}_t = 0$):

$$\boxed{u = \frac{\lambda}{\lambda + \theta q(\theta)}}$$

Esta relación es la **curva de Beveridge**, que define todos los puntos (u,v) para los cuales el desempleo no está creciendo ni decreciendo. Gráficamente:



La intución detrás de que la pendiente sea negativa es que un aumento en las vacantes genera menos estrechez del mercado laboral (mayor θ), la tasa a la que los desempleados encuentran empleo aumenta $(\theta q(\theta))$ y se crean más empleos.

Por otro lado, mientras mayor sea λ, más separaciones exógenas ocurren a cualquier nivel de (u,v), entonces a un mismo v, habrá n mayor u. Esto representa la eficiencia de mercado.

13.3. Ecuaciones de Bellman

Diremos que el producto del trabajo es una constante $p > 0$, pero tener vacantes tiene un costo de $c \cdot p > 0$. Cuando el trabajador está desempleado, percibe un flujo de z. Hay libre entrada de firmas en todo momento, de tal forma que lo limitado es la oferta laboral.

13.3.1. Firmas

V_t será el valor de mantener una vacante y J_t será el valor de llenar un trabajo en t. Entonces,

$$rV_t = -pc + \dot{V}_t + q(\theta_t)\,\text{máx}\,\{J_t - V_t, 0\} \tag{13.1}$$

$$rJ_t = p - w_t + \dot{J}_t + \lambda\,(V_t - J_t) \tag{13.2}$$

Es decir, la firma puede elegir si acepta o rechaza un match (por eso el máx), pero no hay ninguna elección cuando ocurre una separación.

13.3.2. Trabajadores

U_t será el valor de estar desempleado y W_t será el valor de estar empleado. Entonces,

$$rW_t = \dot{W}_t + w_t + \lambda\,(U_t - W_t) \tag{13.3}$$

$$rU_t = \dot{U}_t + z + \theta_t q(\theta_t)\,\text{máx}\,\{W_t - U_t, 0\} \tag{13.4}$$

13.3.3. Equilibrio

Las rentas del match (que se reparten entre trabajadores y firmas) son

$$S_t = (W_t - U_t) + (J_t - V_t)$$

Asumiremos la solución de negociación de Nash (*Generalized Nash Bargaining solution*), donde β representa el poder de negociación de los trabajadores. Entonces,

$$W_t - U_t = \beta S_t, \quad J_t - V_t = (1-\beta)S_t \tag{13.5}$$

Es decir, una fracción β de las rentas del match van al trabajador y otra fracción $1-\beta$ a la firma. Como lo anterior se sostiene para todo t, tenemos que

$$(1-\beta)(\dot{W}_t - \dot{U}_t) = \beta(\dot{J}_t - \dot{V}_t) \tag{13.6}$$

Combinando las ecuaciones (13.1)-(13.6) podemos resolver que el salario será

$$w_t = \beta p + \beta pc + (1-\beta)z + \beta q(\theta_t)(\theta_t - 1)(J_t - V_t) \tag{13.7}$$

Además, la condición de libre entrada nos dice que

$$V_t = 0$$

sustituyendo esto en (13.1) obtenemos que

$$J_t = \frac{pc}{q(\theta_t)}$$

Finalmente, sustituyendo esto en (13.7), llegamos a que

$$w_t = (1-\beta)z + \beta p(1 + c\theta_t) = (1-\beta)z + \beta p + \beta pc\theta_t \tag{13.8}$$

Notemos que el término $(1-\beta)z + \beta p = z + \beta(p-z)$ representa la suma del salario de reserva y la renta que obtiene el trabajador al compartir la renta que genera un trabajo ocupado.

Como $pc\theta_t = pc\frac{v}{u}$, tenemos que este término es el costo promedio de contratación por trabajador desempleado. Entonces $\beta pc\theta_t$ nos dice que el trabajador se lleva una fracción β del ahorro de la firma en costos de contratación cuando se llena la vacante.

Por último, sustituyendo (13.8) en (13.1) se puede mostrar que

$$J_t = \frac{p - w_t}{r + \lambda} = \frac{pc}{q(\theta)} \tag{13.9}$$

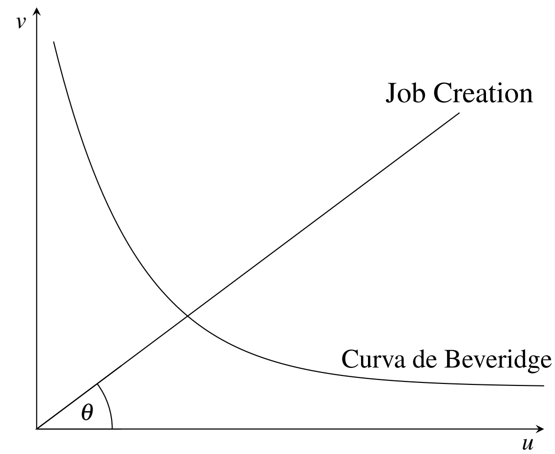
Es decir, en equilibrio el beneficio de llenar una vacante debe ser igual al costo de tener la vacante.

13.4. Estado estacionario

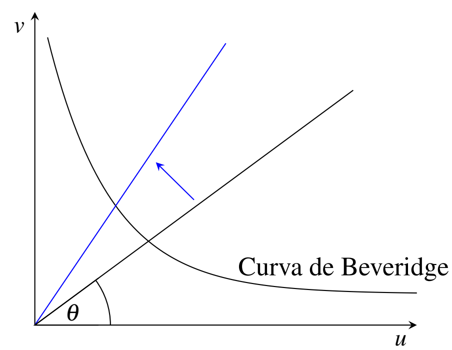
El estado estacionario será el (u, θ, w) que satisfaga

$$\frac{pc}{q(\theta)} = \frac{p - (1-\beta)z + \beta p(1 + c\theta)}{r + \lambda} \tag{13.10}$$

Que es una ecuación resuelta por un único θ^*, que podemos sustituir para obtener $v^* = \theta^* u^*$. Gráficamente

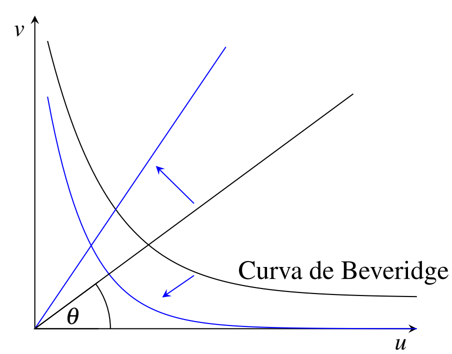


13.4.1. Shock de productividad no anticipado



Un aumento en la productividad ($\uparrow p$) mejora el producto por trabajador, esto hace que el desempleo de equilibrio disminuya ($\downarrow u^*$) y aumenten las vacantes ($\uparrow v^*$).

13.4.2. Shock de separación no anticipado



Ante una disminución de la tasa de separación ($\uparrow \lambda$), junto con un aumento de la productividad ($\uparrow p$), la tasa de desempleo cae inequívocamente ($\downarrow u^*$) y el efecto en la tasa de vacantes es ambiguo.