EC4301 MACROECONOMETRÍA
Asist: Juan Andrés Montero Zúñiga

Laboratorio #2:

Mínimos cuadrados indirectos y prueba de Hausman

Creado: Agosto 2024

Mínimos cuadrados indirectos

Se quiere estimar el siguiente modelo:

$$M_t = a_0 + a_1 Y_t + u_{1t} \tag{1}$$

$$Y_t = b_0 + b_1 M_t + b_2 I_t + u_{2t} \tag{2}$$

Donde:

M_t es la oferta de dinero.

Y_t es la producción.

I_t es la inversión.

u_{1t} y u_{2t} son errores de estimación.

- M es el número de variables endógenas en el modelo; 2
- m es el número de variables endógenas en cada ecuación

- K es el número de variables predeterminadas en el modelo; 1
- k es el número de variables predeterminadas en cada ecuación

> **Importante:**
> Una ecuación está exactamente identificada cuando $K - k = m - 1$.
> Si $K - k > m - 1$ la ecuación está sobre-identificada.
> Si $K - k < m - 1$ la ecuación está sub-identificada.

La ecuación (1) está exactamente identificada porque

$$K - k = 1 - 0 \quad = \quad 2 - 1 = m - 1$$

Sin embargo, la ecuación (2) está sub-identificada porque

$$K - k = 1 - 1 \quad < \quad 2 - 1 = m - 1$$

Note que si no se deja ninguna variable por fuera, la ecuación se vuelve más compleja y como está sub-identificada no podremos saber si los movimientos en las series de tiempo los explica una ecuación o la otra.

Excluir variables predeterminadas de una ecuación (y no de otras) nos permite distinguir situaciones en las que esa ecuación **no se mueve**.

De este modo, nuestra intención es estimar (1) mediante MCI. Para ello debemos encontrar primero el modelo en forma reducida. Expresemos las ecuaciones (1) y (2) como sigue

$$1 \cdot M_t + (-a_1) \cdot Y_t + (-a_0) \cdot 1 + 0 \cdot I_t = u_{1t} \tag{1'}$$
$$(-b_1) \cdot M_t + 1 \cdot Y_t + (-b_0) \cdot 1 + (-b_2) \cdot I_t = u_{2t} \tag{2'}$$

Por lo que en forma matricial tenemos:

$$\underbrace{\begin{pmatrix} 1 & -a_1 \\ -b_1 & 1 \end{pmatrix}}_{\Gamma'} \begin{pmatrix} M_t \\ Y_t \end{pmatrix} + \underbrace{\begin{pmatrix} -a_0 & 0 \\ -b_0 & -b_2 \end{pmatrix}}_{B'} \begin{pmatrix} 1 \\ I_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

Pre-multiplicamos ambos lados de la igualdad por $(\Gamma')^{-1}$, donde:

$$(\Gamma')^{-1} = \frac{1}{1 - a_1 b_1} \begin{pmatrix} 1 & a_1 \\ b_1 & 1 \end{pmatrix}$$

Así que:

$$\begin{pmatrix} M_t \\ Y_t \end{pmatrix} = \frac{1}{1 - a_1 b_1} \begin{pmatrix} 1 & a_1 \\ b_1 & 1 \end{pmatrix} \begin{pmatrix} a_0 & 0 \\ b_0 & b_2 \end{pmatrix} \begin{pmatrix} 1 \\ I_t \end{pmatrix} + \frac{1}{1 - a_1 b_1} \begin{pmatrix} 1 & a_1 \\ b_1 & 1 \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

$$\begin{pmatrix} M_t \\ Y_t \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{a_0 + a_1 b_0}{1 - a_1 b_1} & \frac{a_1 b_2}{1 - a_1 b_1} \\ \frac{b_1 a_0 + b_0}{1 - a_1 b_1} & \frac{b_2}{1 - a_1 b_1} \end{pmatrix}}_{\Pi'} \begin{pmatrix} 1 \\ I_t \end{pmatrix} + \underbrace{\begin{pmatrix} \frac{u_{1t} + a_1 u_{2t}}{1 - a_1 b_1} \\ \frac{b_1 u_{1t} + u_{2t}}{1 - a_1 b_1} \end{pmatrix}}_{v_t} \tag{6}$$

Entonces, la forma reducida del modelo es:

$$M_t=\frac{a_0+a_1b_0}{1-a_1b_1}+\frac{a_1b_2}{1-a_1b_1}I_t+\frac{u_{1t}+a_1u_{2t}}{1-a_1b_1} \tag{7}$$

$$Y_t=\frac{b_1a_0+b_0}{1-a_1b_1}+\frac{b_2}{1-a_1b_1}I_t+\frac{b_1u_{1t}+u_{2t}}{1-a_1b_1} \tag{8}$$

La cual podemos escribir como:

$$M_t=\pi_0+\pi_1I_t+v_{1t} \tag{7'}$$

$$Y_t=\pi_2+\pi_3I_t+v_{2t} \tag{8'}$$

Donde:

$$\pi_0:=\frac{a_0+a_1b_0}{1-a_1b_1},\quad \pi_1:=\frac{a_1b_2}{1-a_1b_1},\quad \pi_2:=\frac{b_1a_0+b_0}{1-a_1b_1},\quad \pi_3:=\frac{b_2}{1-a_1b_1}$$

y los choques reducidos (en función de los estructurales) son:

$$v_{1t}:=\frac{u_{1t}+a_1u_{2t}}{1-a_1b_1},\quad v_{2t}:=\frac{b_1u_{1t}+u_{2t}}{1-a_1b_1}$$

Sabemos que los parámetros estructurales que podemos estimar son a_0 y a_1; así que obtendremos sus valores en términos de los parámetros del modelo reducido:

$$\frac{\pi_1}{\pi_3}=\frac{\frac{a_1b_2}{1-a_1b_1}}{\frac{b_2}{1-a_1b_1}}=a_1 \tag{9}$$

$$\pi_0-\frac{\pi_1\pi_2}{\pi_3}=\frac{a_0+a_1b_0}{1-a_1b_1}-a_1\frac{b_1a_0+b_0}{1-a_1b_1}=\frac{a_0+a_1b_0-a_1b_1a_0+a_1b_0}{1-a_1b_1}=a_0 \tag{10}$$

Provistos de estas fórmulas, nuestro trabajo en términos de escritura de código es diminuto: solo debemos estimar (7') y (8') mediante mínimos cuadrados ordinarios (MCO).

Estimación en Stata

Ya simplificados los coeficientes en Stata será muy simple obtenerlos, haremos una simple regresión de la forma siguiente:

```
regress m i
```

```
regress y i
```

Se obtiene

$$\hat{\pi}_0\approx 312.0608$$

$$\hat{\pi}_1\approx 0.5692502$$

$$\hat{\pi}_2\approx 852.3203$$

$$\hat{\pi}_3\approx 5.352226$$

Sustituyendo los valores en la fórmula anterior:

$$\hat{a}_0^{MCI} \approx 312.0608 - \frac{(0.5692502)(852.3203)}{5.352226} \approx 221.410013 \quad (11)$$

$$\hat{a}_1^{MCI} \approx \frac{0.5692502}{5.352226} \approx 0.10636 \quad (12)$$

Comparemos esto con $\hat{a}_0^{MCO}$ y $\hat{a}_1^{MCO}$ corriendo el siguiente código,

```
regress m y
```

Se obtiene:

$$\hat{a}_0^{MCO} \approx 162.70$$

$$\hat{a}_1^{MCO} \approx 0.1159$$

Los resultados difieren a simple vista: por desgracia, es difícil superar esta comparación informal ya que la alta volatilidad que caracteriza a los estimadores de MCI generalmente impide corroborar si las diferencias son estadísticamente significativas.

No obstante, una posible explicación para esas diferencias es la existencia de sesgo de simultaneidad; y esta sí que podemos evaluarla con una prueba (de Hausman)

> **Sesgo de simultaneidad**: Es donde la variable explicativa se determina conjuntamente con la variable dependiente . En otras palabras, X causa Y pero Y también causa X.

Fuente: statologos

Prueba de Hausman

1. Se estima la ecuación sub-identificada con una regresión `regress y i`
2. Guarda los residuos de la regresión `predict y_res, res`
3. Guarda la variable endógena estimada. `predict y_hat, xb`
4. Estima la ecuación exactamente identificada, reemplazando la variable endógena por la estimada e incluyendo sus choques reducidos. `regress m y_hat y_res`

Note que `y_res` es significativo para todos los niveles de significancia convencionales, es decir, **se rechaza** la hipótesis de que el coeficiente asociado a $\hat{v}_{1t}$ sea nulo: evidencia que apunta a la presencia de simultaneidad y por ende un sesgo (que no desaparece asintóticamente) en los estimadores de MCO y por lo tanto lo correcto es la estimación con MCI.

> Nota: Los estimadores MCI heredan la **consistencia** de los MCO.

Modelo de Oferta y Demanda con Precio Endógeno

Consideremos un mercado donde el precio del bien (P) es determinado dentro del sistema por las interacciones entre la oferta (Q_s) y la demanda (Q_d) del bien.

Ecuaciones del Modelo:

1. **Ecuación de Demanda**:

$$Q_d = \alpha_0 + \alpha_1 P + \alpha_2 I + u_1$$

Donde:

- Q_d es la cantidad demandada.
- P es el precio del bien (variable endógena).
- I es el ingreso de los consumidores (variable exógena).
- u_1 es el término de error.

2. **Ecuación de Oferta**:

$$Q_s = \beta_0 + \beta_1 P + \beta_2 C + u_2$$

Donde:

- Q_s es la cantidad ofrecida.
- P es el precio del bien (variable endógena).
- C es el costo de producción (variable exógena).
- u_2 es el término de error.

1. **Ecuación de Equilibrio de Precio**:

$$P = \gamma_0 + \gamma_1 I + \gamma_2 C + u_3$$

Donde:
- P es el precio del bien (variable endógena).
- I es el ingreso de los consumidores (variable exógena).
- C es el costo de producción (variable exógena).
- u_3 es el término de error.

Despeje y Modelo Reducido

$$Q_d = \alpha_0 + \alpha_1(\gamma_0 + \gamma_1 I + \gamma_2 C + u_3) + \alpha_2 I + u_1$$

$$Q_s = \beta_0 + \beta_1(\gamma_0 + \gamma_1 I + \gamma_2 C + u_3) + \beta_2 C + u_2$$

Agrupando:

$$Q_d = (\alpha_0 + \alpha_1\gamma_0) + (\alpha_1\gamma_1 + \alpha_2)I + \alpha_1\gamma_2 C + (\alpha_1 u_3 + u_1)$$

$$Q_s = (\beta_0 + \beta_1\gamma_0) + \beta_1\gamma_1 I + (\beta_1\gamma_2 + \beta_2)C + (\beta_1 u_3 + u_2)$$

Forma Reducida

- **Demanda Reducida**:

$$Q_d = \pi_0 + \pi_1 I + \pi_2 C + v_1$$

- **Oferta Reducida**:

$$Q_s = \pi_3 + \pi_4 I + \pi_5 C + v_2$$

Donde:

$$\pi_0 = \alpha_0 + \alpha_1\gamma_0$$

$$\pi_1 = \alpha_1\gamma_1 + \alpha_2$$

$$\pi_2 = \alpha_1\gamma_2$$

$$v_1 = \alpha_1 u_3 + u_1$$

$$\pi_3 = \beta_0 + \beta_1\gamma_0$$

$$\pi_4 = \beta_1\gamma_1$$

$$\pi_5 = \beta_1\gamma_2 + \beta_2$$

$$v_2 = \beta_1 u_3 + u_2$$

Generación de Datos en Stata

Se generan los datos en Stata de la siguiente forma:

```
clear all
set seed 12345
set obs 100

* Generar variables exógenas
gen I = rnormal(300, 20)   // Ingreso de los consumidores
gen C = rnormal(100, 30)   // Costo de producción

* Generar el precio (endógeno) usando una ecuación de equilibrio con un término
de error
gen P = 5 + 0.1*I + 0.2*C + rnormal(0, 1)

* Generar las cantidades demandadas y ofrecidas basadas en el precio
gen Q_d = 200 - 5*P + 0.1*I + rnormal(0, 1)
gen Q_s = -50 + 4*P - 0.3*C + rnormal(0, 2)
```

Note los parámetros de P puedes estimarse directamente.

`regress P I C`

Se obtiene:

$$\hat{\gamma_0} \approx 0.0914621$$

$$\hat{\gamma_1} \approx 0.2038071$$

$$\hat{\gamma_2} \approx 7.408306$$

Estimación de MCI en Stata

`regress Q_d I C`

`regress Q_s I C`

Se obtienen los parámetros reducidos:

$$\hat{\pi}_0 \approx 163.5419$$

$$\hat{\pi}_1 \approx -.3587771$$

$$\hat{\pi}_2 \approx -1.021651$$

$$\hat{\pi}_3 \approx -17.04442$$

$$\hat{\pi}_4 \approx .3584955$$

$$\hat{\pi}_4 \approx .5053602$$

Y listo, ahora únicamente debe despejar los α y β.

Prueba de Hausman

```
regress Q_d I C
regress Q_s I C
* Guardar los resultados de MCO
estimates store mco
gen P_hat_d = (_b[I] * I + _b[C] * C)  // Precio estimado a partir de la
ecuación de demanda reducida
gen P_hat_s = (_b[I] * I + _b[C] * C)  // Precio estimado a partir de la
ecuación de oferta reducida

reg Q_d P_hat_d I  // Regresión de la ecuación de demanda con P_hat_d
reg Q_s P_hat_s C  // Regresión de la ecuación de oferta con P_hat_s
estimates store mci


* Realizar la prueba de Hausman
hausman mci mco

*Dado que el valor de p es extremadamente bajo (0.0000), rechazamos la
hipótesis nula. Esto significa que existe evidencia estadísticamente
significativa de que las estimaciones por MCO están sesgadas.
```