PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
CIENCIAS SOCIALES
CICLO 2022-2

Fundamentos de Econometría
<u>Práctica Dirigida 10</u>

Profesor:	Juan Palomino	juan.palominoh@pucp.pe
Jefes de Práctica:	Tania Paredes	tania.paredes@pucp.edu.pe

Fecha: 26 – 11 – 2022

---

1. Endogeneidad

Dado el siguiente modelo de determinación de la oferta monetaria:

$$Y_{1t} = \beta_{10} + \beta_{11} Y_{2t} + \gamma_{11} X_{1t} + \gamma_{12} X_{2t} + u_{1t} \qquad \text{(Ingreso)}$$

$$Y_{2t} = \beta_{20} + \beta_{21} Y_{1t} + u_{2t} \qquad \text{(Oferta Monetaria)}$$

En donde Y_{1t} corresponde al ingreso de los hogares, Y_{2t} a la oferta monetaria, X_{1t} al gasto de inversión y X_{2t} al gasto público. u_{1t} y u_{2t} son variables aleatorias gausianas bien comportadas.

Estimaremos la ecuación de la oferta monetaria utilizando mínimos cuadrados en 2 etapas (MC2E).

Primera etapa

a. Plantear la forma reducida del modelo. ¿En qué se diferencia de la forma estructural?

$$Y_{1t} = \beth_0 + \beth_1 X_{1t} + \beth_2 X_{2t} + u_t$$
$$Y_{2t} = \beth_3 + \beth_4 X_{1t} + \beth_5 X_{2t} + u_t$$

b. Estudiar la identificación por medio de la condición de orden y de rango.

Condición de orden:

La ecuación 1 (Y_1t) no excluye ninguna variable (K-kj=0), e incluye 1 a la derecha de la ecuación (gj-1=1). Dado que 0<1, la ecuación está subidentificada.
La ecuación 2 (Y_2t) excluye dos variables exógenas (K-kj=2) y no incluye a ninguna variable (gj-1=0). Por lo tanto está sobreidentificada.

c. Estimar la ecuación del ingreso en forma reducida por medio de MCO.

$$(\hat{\beth}_0, \hat{\beth}_1, \hat{\beth}_2) = inv(X_a{}' X_a) X_a{}' Y_a$$

Donde: $X_a = (\mathbf{1}, \boldsymbol{X_1}, \boldsymbol{X_2})$ y $Y_a = (\boldsymbol{Y_1})$

d. Expresar $\hat{Y}_{1t}$ como una combinación lineal de las variables exógenas X_{1t} y X_{2t} .

$$\hat{Y}_{1t} = \hat{\beth}_0 + \hat{\beth}_1 X_{1t} + \hat{\beth}_2 X_{2t}$$

Segunda etapa

e. Introducir $Y_{1t} = \hat{Y}_{1t} + u_{1t}$ en la ecuación estructural de la oferta monetaria.

$$\begin{aligned} Y_{2t} &= \beta_{20} + \beta_{21}(\hat{Y}_{1t} + \hat{u}_t) + u_{2t} \\ &= \beta_{20} + \beta_{21}\hat{Y}_{1t} + (u_{2t} + \beta_{21}\hat{u}_t) \\ &= \beta_{20} + \beta_{21}\hat{Y}_{1t} + u_t^* \end{aligned} \qquad (20.4.6)$$

en donde $u_t^* = u_{2t} + \beta_{21}\hat{u}_t$.

f. Estimar la ecuación estructural de la forma monetaria por MCO

$$(\beta_{20}, \beta_{21}) = inv(\hat{Y}_a{}'\hat{Y}_a)\hat{Y}_a{}'Y_b$$

Donde: $\hat{Y}_a = (\mathbf{1}, \boldsymbol{\hat{Y}_1})$ y $Y_b = (\boldsymbol{Y_2})$

2. Modelo Ecuaciones Simultáneas

Considere el siguiente modelo de demanda y oferta de dinero:

Demanda de dinero: $M_t^d = \beta_0 + \beta_1 Y_t + \beta_2 R_t + \beta_3 P_t + u_{1t}$

Oferta de dinero: $M_t^s = \alpha_0 + \alpha_1 Y_t + u_{2t}$

En donde: M = dinero, Y = ingreso, R = tasa de interés, P = precio. Suponga que R y P están predeterminados.

- Identificamos las variables del sistema:
  - Variables endógenas: $\mathrm{M_t}, Y_t$
  - Variables exógenas del sistema: (1), R_t, P_t

- La condición de orden requiere conocer cómo está compuesto el modelo:
  - G: Número de endógenas del sistema
  - gj: Número de endógenas de la ecuación
  - K: Número de exógenas del sistema
  - kj: Número de exógenas de la ecuación

- Por otro lado, del siguiente criterio: **en el cual las exógenas excluidas = endógenas incluidas**

$$\boldsymbol{K - k_j \geq g_j - 1}$$

a. ¿Está identificada la función de demanda? Justifique.

$$\mathbf{3 - 3 \geq 2 - 1}$$
$$\mathbf{0 < 1}$$

El modelo está subidentificado. No será posible obtener una solución a partir de los datos.

b. ¿Está identificada la función de oferta? Justifique.

$$\mathbf{3 - 1 \geq 2 - 1}$$
$$\mathbf{2 > 1}$$

El modelo está sobreidentificado. No será posible obtener una solución a partir de los datos.

c. ¿Cuál método se utilizaría para estimar los parámetros de la(s) ecuación(es) identificada(s)? ¿Por qué?

En el caso que hemos analizado, no encontramos que ninguna ecuación esté identificada. No obstante, aún podemos analizar el caso de la función de oferta dado que esta está sobreidentificada. Para realizar ello, podemos utilizar el método de Mínimos Cuadrados en 2 Etapas (MC2E).

d. Suponga que se modifica la función de oferta agregando las variables explicativas Y_{t-1} y M_{t-1}. ¿Qué sucede con el problema de la identificación? ¿Se utilizaría aún el método que utilizó en (c)? ¿Por qué sí o por qué no?

- Identificamos las variables del sistema:
  - Variables endógenas: M_t, Y_t
  - Variables exógenas del sistema: (1), R_t, P_t, Y_{t-1}, M_{t-1}

  Analizamos el caso de la demanda:

$$\mathbf{5 - 3 \geq 2 - 1}$$
$$\mathbf{2 > 1}$$

  Analizamos el caso de la oferta:

$$\mathbf{5 - 3 \geq 2 - 1}$$
$$\mathbf{2 > 1}$$

La ecuación de la demanda esta aún subidentificada y la oferta aún siguen estando sobreidentificadas. Por lo que, el uso de MC2E seguirá siendo el ideal para resolver este modelo de la oferta como se ha planteado en c.

3. Modelo de Variable Dependiente Binaria

a. La siguiente estimación presenta factores asociados a que un niño o niña sufra desnutrición. Identifique e interprete los siguientes aspectos: el ajuste conjunto, el ajuste individual y los coeficientes estimados.

```
. logit des_cro urbana sexo2 v012 v133 peso_n leng_indi bord wi v136

Iteration 0:   log likelihood = -2298.2328
Iteration 1:   log likelihood =  -2027.551
Iteration 2:   log likelihood = -1972.0705
Iteration 3:   log likelihood = -1971.7715
Iteration 4:   log likelihood = -1971.7714

Logistic regression                               Number of obs   =       7111
                                                  LR chi2(9)      =     652.92
                                                  Prob > chi2     =     0.0000
Log likelihood = -1971.7714                       Pseudo R2       =     0.1420
```

des_cro	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
urbana	-.1311455	.1143819	-1.15	0.252	-.3553298	.0930389
sexo2	-.1793039	.0847413	-2.12	0.034	-.3453938	-.013214
v012	-.0317248	.0086172	-3.68	0.000	-.0486142	-.0148354
v133	-.0374103	.0144909	-2.58	0.010	-.065812	-.0090087
peso_n	-1.029762	.0776204	-13.27	0.000	-1.181895	-.8776285
leng_indi	-.1525679	.143083	-1.07	0.286	-.4330054	.1278697
bord	.170544	.0383674	4.45	0.000	.0953452	.2457427
wi	-.6051935	.0753233	-8.03	0.000	-.7528246	-.4575625
v136	.0357305	.0227069	1.57	0.116	-.0087742	.0802351
_cons	1.741954	.3812959	4.57	0.000	.9946283	2.489281

Solución:

En principio, recordar de la PD8 que los modelos logit y probit son no lineales, por lo que estos no podrán estimarse por MCO o Mínimos Cuadrados Ponderados (MCP) como en el caso del Modelo de Probabilidad Lineal (MPL), por lo que este tipo de modelos se estima por Máxima Verosimilitud. Como la estimación de máxima verosimilitud esta basada en la distribución de *y* dado *x*, la heterocedasticidad en Var(y|x) automáticamente se toma en cuenta.

Respecto al output de la estimación, esta se obtiene del proceso iterativo de la estimación por MV:

- LR ch2(9) será el simil del f-estadístico en un modelo de regresión lineal y que tenga un Prob>chi2 al 0.000, implica que el modelo es significativo conjuntamente al 5%.
- El ajuste conjunto se obtiene del Pseudo-R cuadrado, el cual compara el modelo log-likehood del modelo estimado y el modelo nulo (el cual no incluye ninguna de las variables). <u>OJO: Esta variable no representa la</u>

<u>proporción de la varianza explicada por el modelo.</u> Ello determina si el modelo está siendo explicado adecuadamente por las variables, si este tiene un valor cercano a 1, implica que el modelo se ajusta perfectamente. Un pseudo R cuadrado entre 0.2 y 0.4 es el ideal. En este caso, tenemos un pseudo R cuadrado mayor a cero, por lo que las variables estimadas están explicando de alguna forma el modelo.

- El ajuste individual se obtiene p-value e intervalo de confianza para cada variable. Por ejemplo, las variables bord (número de orden de nacimiento), wi (índice de nivel socioeconómico) y v133 (años de educación de la madre) son significativos al observar el p-value.
- Los coeficientes estimados reportan el signo del impacto y, para conocer su nivel de significancia al 5%, se revisará el p-value Por ejemplo, la variable bord tendrá un efecto positivo sobre la desnutrición y su efecto será significativo al 5%. En este caso, los valores obtenidos de los coeficientes no indicarán la magnitud de impacto.

b. Compare los resultados de la tabla en 1a (de estimación logit) con los resultados obtenidos de una estimación probit con las mismas variables. ¿Qué tanto varía el ajuste individual, ajuste conjunto y los coeficientes?

```
. probit des_cro urbana sexo2 v012 v133 peso_n leng_indi bord wi v136

Iteration 0:   log likelihood = -2298.2328
Iteration 1:   log likelihood = -1983.3704
Iteration 2:   log likelihood = -1970.2702
Iteration 3:   log likelihood = -1970.2208
Iteration 4:   log likelihood = -1970.2208

Probit regression                               Number of obs   =       7111
                                                LR chi2(9)      =     656.02
                                                Prob > chi2     =     0.0000
Log likelihood = -1970.2208                     Pseudo R2       =     0.1427
```

des_cro	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
urbana	-.0812933	.0607533	-1.34	0.181	-.2003677	.0377811
sexo2	-.0918056	.0444721	-2.06	0.039	-.1789694	-.0046418
v012	-.0169358	.0044503	-3.81	0.000	-.0256582	-.0082134
v133	-.0186873	.0076192	-2.45	0.014	-.0336206	-.003754
peso_n	-.5478043	.0409281	-13.38	0.000	-.6280219	-.4675866
leng_indi	-.0939845	.0739025	-1.27	0.203	-.2388308	.0508619
bord	.0981412	.020758	4.73	0.000	.0574563	.138826
wi	-.3037867	.0380512	-7.98	0.000	-.3783657	-.2292078
v136	.0184192	.0121241	1.52	0.129	-.0053437	.0421821
_cons	.8301597	.2004283	4.14	0.000	.4373275	1.222992

Solución:

Respecto al output de la estimación, esta se obtiene del proceso iterativo de la estimación por MV:

- LR ch2(9) será el simil del f-estadístico en un modelo de regresión lineal y que tenga un Prob>chi2 al 0.000, implica que el modelo es significativo conjuntamente al 5%.

- El ajuste conjunto se obtiene del Pseudo-R cuadrado, el cual compara el modelo log-likehood del modelo estimado y el modelo nulo (el cual no incluye ninguna de las variables). Ello determina si el modelo está siendo explicando adecuadamente por las variables, si este tiene un valor cercano a 1, implica que el modelo se ajusta perfectamente. Un pseudo R cuadrado entre 0.2 y 0.4 es el ideal. En este caso, tenemos un pseudo R cuadrado mayor a cero, por lo que las variables están explicando el modelo de alguna forma.
- Los coeficientes estimados reportan el signo del impacto y, para conocer su nivel de significancia al 5%, se revisará el p-value Por ejemplo, la variable bord tendrá un efecto positivo sobre la desnutrición y su efecto será significativo al 5%. En este caso, los valores obtenidos de los coeficientes no indicarán la magnitud de impacto. En este caso, los valores obtenidos de los coeficientes no indicarán la magnitud de impacto.
- Si se compara las otras variables en torno a los signos y p-value, estos tendrán un comportamiento igual que en modelo estimado por logit.

4. Series de Tiempo

a. Escriba los dos primeros momentos de una serie y defina qué es estacionariedad débil (llamado a veces simplemente estacionariedad) y estacionariedad fuerte.

Solución:

Los dos primeros momentos de una serie son los siguientes:

Esperanza: $E[Y_t] = \mu < \infty \quad , \forall t$

Covarianza: $E[(Y_t - \mu)(Y_{t-j} - \mu)] = \gamma_j < \infty \quad , \forall t, j$

Donde la autocovarianza es la varianza:

$$E[(Y_t - \mu)(Y_t - \mu)] = E[(Y_t - \mu)^2] = Var(Y_t) = \gamma_0$$

Una serie es estacionaria si sus dos primeros momentos no dependen del tiempo. Por otro lado, una serie tiene estacionariedad fuerte si todos sus momentos no dependen del tiempo (por ejemplo, el tercer momento es la asimetría y el cuarto momento es la kurtosis)

b. Defina que es un ruido blanco (*white noise* en inglés) y mencione sus supuestos

Solución:

Un ruido blanco ε_t es un proceso con valores independientes e idénticamente distribuidos que no dependen del tiempo y cumplen las siguientes condiciones:

$E(\varepsilon_t) = 0 \quad \forall t$

$E(\varepsilon_t^2) = \sigma^2 \quad \forall t$

$E[\varepsilon_t \varepsilon_j] = 0 \quad \forall t \neq j$

c. Exprese los siguientes procesos de series de tiempo:

   4.c.1. AR(1)
   4.c.2. MA(1)
   4.c.3. AR(p)
   4.c.4. MA(q)
   4.c.5. ARMA(p,q)

Solución:

Considerando procesos sin deriva (llamado pendiente o drift en inglés).

i. AR(1) : $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$

iii. MA(1) : $Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$

ii. AR(p) : $Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$

iv. MA(q) : $Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$

v. ARMA(p,q):

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

d. Simplifique las expresiones de 1c usando el operador de rezagos y defina el polinomio de rezagos y el polinomio característico.

Solución:

i. AR(1) : $Y_t = \phi_1 L Y_t + e_t$

$$(1 - \phi_1 L) Y_t = \varepsilon_t$$

iii. MA(1) : $Y_t = \varepsilon_t + \theta_1 L \varepsilon_t$

$$Y_t = (1 + \theta_1 L) \varepsilon_t$$

ii. AR(P) : $Y_t = \phi_1 L Y_t + \dots + \phi_p L^p Y_t + \varepsilon_t$

$$(1 - \phi_1 L - \dots - \phi_p L^p) Y_t = \varepsilon_t$$

iv. MA(q) : $Y_t = \varepsilon_t + \theta_1 L e_t + \dots + \theta_q L^q e_t$

$$Y_t = (1 + \theta_1 L + \dots + \theta_q L^q) \varepsilon_t$$

v. ARMA(p,q):

$$Y_t = \phi_1 L Y_t + \dots + \phi_p L^p Y_t + \theta_1 L e_t + \dots + \theta_q L^q e_t + e_t$$

$$(1 - \phi_1 L - \dots - \phi_p L^p) Y_t = (1 + \theta_1 L + \dots + \theta_q L^q) e_t$$

El polinomio de rezagos son los polinomios factorizados de los procesos de series de tiempo.

Los polinomios de rezagos descritos son los siguientes:

i. AR(1) : $\Phi_1(L) = (1 - \phi_1 L)$

iii. MA(1) : $\Theta_1(L) = (1 + \theta_1 L)$

ii. AR(P) : $\Phi_p(L) = (1 - \phi_1 L - \dots - \phi_p L^p)$

iv. MA(q) : $\Theta_q(L) = (1 + \theta_1 L + \dots + \theta_p L^p)$

v. ARMA(p,q): $\Phi_p(L) = (1 - \phi_1 L - \dots - \phi_p L^p)$ y

$$\Theta_q(L) = (1 + \theta_1 L + \dots + \theta_p L^p)$$

Los polinomios característicos transforma al polinomio de rezagos en las siguientes expresiones:

i. AR(1) : $\Phi_1(\lambda) = (\lambda - \phi_1)$

iii. MA(1) : $\Theta_1(L) = (\lambda + \theta_1)$

ii. AR(P) : $\Phi_p(L) = (\lambda^p - \phi_1 \lambda^{p-1} \dots - \phi_p)$

iv. MA(q) : $\Theta_q(L) = (\lambda^q + \theta_1 \lambda^{q-1} + \dots + \theta_p)$

v. ARMA(p,q): $\Phi_p(L) = (\lambda^p - \phi_1 \lambda^{p-1} \dots - \phi_p)$ y

$$\Theta_q(L) = (\lambda^q + \theta_1 \lambda^{q-1} + \dots + \theta_p)$$

e. Muestre que puede convertir un AR(1) en un MA(∞) ¿Qué condición debe de cumplir el parámetro de persistencia del AR(1)? Asocie a ello con el concepto de estacionariedad e invertibilidad.

Solución:

El AR(1) con operador de rezagos es el siguiente:

$$(1-\phi_1 L)Y_t = \varepsilon_t$$

$$Y_t = \left(\frac{1}{1-\phi_1 L}\right)\varepsilon_t$$

$$Y_t = (1+\phi_1 L+\phi_1^2 L^2+\dots)\varepsilon_t$$

$$Y_t = \varepsilon_t+\phi_1\varepsilon_{t-1}+\phi_1^2\varepsilon_{t-2}+\dots$$

$$Y_t = \sum_{i=0}^{\infty}\phi_1^i\varepsilon_{t-i}$$

Ello es un MA(∞). La condición de invertibilidad permite invertir un modelo AR(p) en un MA(∞).

El AR(1) será estacionario si $|\phi_1|<1$.

f. ¿Qué condición debe de cumplir el polinomio de rezagos y el polinomio característico para que la serie sea estacionaria?

Solución:

Para ello, analizamos ambos polinomios. Para que una serie sea estacionaria, necesariamente $|\phi_1| < 1$. Si analizamos el la raíz del polinomio de rezagos:

$$\phi_1(L) = 1 - \phi_1 L = 0$$

$$|L| = \left|\frac{1}{\phi_1}\right|,$$

donde si $|\phi_1| < 1$, entonces la raíz del polinomio de rezagos debe estar fuera del círculo unitario.

Si analizamos la raíz del polinomio característico:

$$\phi_1(\lambda) = \lambda_1 - \phi_1 = 0$$

$$|\lambda_1| = |\phi_1|$$

donde si $|\phi_1| < 1$, entonces la raíz del polinomio característico debe estar dentro del círculo unitario.

Lo mismo se cumplirá para los procesos AR(p).

g. Repita los procesos descritos en 1e y 1f para el proceso MA(1) y defina el concepto de invertibilidad y su implicancia con los polinomios descritos.

Solución:

Existe el concepto de invertibilidad para un MA(1).

$$MA(1): \quad Y_t = (1+\theta_1 L)\varepsilon_t$$
$$Y_t = [1 - \underbrace{(-\theta_1)}L]\varepsilon_t$$
$$Y_t = [1 - \varphi L]\varepsilon_t$$
$$\frac{1}{[1-\varphi L]} Y_t = \varepsilon_t$$
$$(1 + \varphi L + \varphi^2 L^2 + \dots) Y_t = \varepsilon_t$$
$$Y_t = -\varphi Y_{t-1} - \varphi^2 Y_{t-2} - \dots + \varepsilon_t$$
$$Y_t = \varepsilon_t - \sum_{i=1}^{\infty} \varphi^i Y_{t-i}$$

Puedo expresar un MA(1) como un AR(∞). Esta es la condición de invertibilidad de un MA. Además, para ello $|\varphi| < 1$; por lo tanto, $|\theta_1| < 1$.

Podemos hallar lo mismo a través de los polinomios descritos anteriormente, enfocandonos en los modelos MA. Para que sea invertible, necesariamente $|\theta_1| < 1$. Si analizamos el la raíz del polinomio de rezagos:

$$\theta_1(L) = 1 + \theta_1 L = 0$$
$$|L| = \left|\frac{1}{\theta_1}\right|,$$

donde si $|\theta_1| < 1$, entonces la raíz del polinomio de rezagos debe estar fuera del círculo unitario.

Si analizamos la raíz del polinomio característico:

$$\theta_1(\lambda) = \lambda_1 - \theta_1 = 0$$
$$|\lambda_1| = |\theta_1|$$

donde si $|\theta_1| < 1$, entonces la raíz del polinomio característico debe estar dentro del círculo unitario.

Lo mismo se cumplirá para los procesos MA(q).

h. Comente las implicancias de estacionariedad e invertibilidad para un ARMA(p,q)

Solución:

En un modelo ARMA(p,q), si analizamos estacionariedad nos enfocamos en la parte AR y si analizamos invertibilidad nos enfocamos en la parte MA. Ello es posible dado que todo MA finito es estacionario (suma finita de ruido blanco). Además, todo proceso AR es invertible (mostrado líneas arriba).

- Un proceso ARMA(p,q) es estacionario sí y solo sí el módulo de las raíces del polinomio de rezagos (característico) del componente autorregresivo está fuera (dentro) del círculo unitario.
- Un proceso ARMA(p,q) es invertible sí y solo sí el módulo de las raíces del polinomio de rezagos (característico) del componente de medias móviles está fuera (dentro) del círculo unitario.