
EP4179 — Microeconometría

Luis Chávez
Universidad Nacional Agraria La Molina

2024-II
Pset3: IV

A. Herramientas básicas

Problema 1. Sea el modelo básico:

$$peso_i = \beta_0 + \beta_1 talla_i + \beta_2 d_i$$

donde d_i es una variable indicadora que representa 1 si la persona i consume KFC y 0 en caso contrario.

- a) Establecer cuál sería el(los) problema(s) de estimar el modelo propuesto por OLS.
- b) ¿Se podría definir un instrumento para d_i , asumiendo que $E(d_i u_i) \neq 0$?

Problema 2. Suponga que se desea medir la incidencia del consumo de cerveza en el número de embarazos (m_i) en 92 universidades peruanas. A partir de una inspección teórica sistemática se ha planteado el modelo transversal:

$$m_i = \beta_0 + \beta_1 c_i + \beta_2 f x_i + \beta_3 ta_i + u_i, \quad \forall i = 1, \dots, 92$$

donde c_i es el consumo promedio de cerveza (en litros), $f x_i$ es una dummy que indica 1 si la universidad realiza festivales extra académicos (0 en caso contrario) y ta es el tamaño del campus universitario. Se pide:

- a) Precisar si se verifica el supuesto de exogeneidad al estimar el modelo vía OLS. Explicar en qué variable está el problema (de ser el caso).
- b) Establecer si el modelo estará correctamente especificado. Señale variables específicas para argumentar.
- c) ¿Qué variables podrían ser instrumentos óptimos de c_i si ahora la variable dependiente es $\ln(st_i)$, donde st_i es la suma de las notas de los 200 estudiantes más destacados? Demostrar su optimalidad.

B. Herramientas intermedias

Problema 3. Cueviña ha dejado el fútbol para estudiar economía. Ha planteado el modelo estructural:

$$y_i = \alpha_1 x_1 + \alpha_2 h_i + \epsilon_i$$

Se sabe que $E(z\epsilon) = 0$, $E(h\epsilon) \neq 0$ y el vector de variables exógenas es $x' = (x_1, x_2)$. Cueviña ha propuesto realizar dos pasos: i) Regresionar h_i sobre x_2 y obtener el predictor \hat{h}_i y ii) Regresionar y_i sobre x_1 y \hat{h}_i , para obtener el vector de estimadores $(\hat{\alpha}_1, \hat{\alpha}_2)'$. Se pide:

- a) Demostrar la consistencia de los estimadores obtenidos.
- b) Ayude a Cueviña a interpretar $\hat{\alpha}_2$ si la variable dependiente está expresado en millones y en logaritmos; además h_i está en metros por cada 50 mil.

C. Herramientas avanzadas

Problema 4. (Adaptado de Cameron & Trivedi, 2005) Sea el modelo de 3 ecuaciones:

$$y_i = \beta x_i + u_i$$

$$x_i = \theta u_i + v_i$$

$$z_i = \gamma v_i + e_i$$

donde todos los términos de error son independientes entre sí y siguen una media 0 y varianza σ_u^2 , σ_v^2 y σ_e^2 , respectivamente. Se pide demostrar:

- a) $plim(\hat{\beta} - \beta) = \theta\sigma_u^2/(\theta\sigma_u^2 + \sigma_v^2)$.
- b) $\rho_{xz} = \gamma\sigma_v^2/(\theta^2\sigma_u^2 + \sigma_v^2)(\gamma^2\sigma_v^2 + \sigma_e^2)$.
- c) $(\hat{\beta}_{IV} - \beta) \rightarrow 1/\theta$ conforme $\gamma \rightarrow 0$