Econometría 1 - Actividad 7

Variables Instrumentales

Ejercicios 1 y 2: Teoría y Derivaciones

División de Economía - CDE
Dr. Francisco Cabrera

1. EJERCICIO 1: Efecto de PC sobre GPA

1.1. Ejercicio 1a: Endogeneidad de PC

El modelo

Consideramos el modelo:

$$\text{GPA}_i = \beta_0 + \beta_1 \text{PC}_i + u_i \tag{1}$$

donde:

- GPA_i = promedio de calificaciones del estudiante i
- PC_i = variable binaria (1 si tiene computadora, 0 si no)
- u_i = término de error (incluye todas las variables no observadas)

El problema de endogeneidad

Supuesto de Exogeneidad (necesario para MCO):

$$E(u_i \mid \text{PC}_i) = 0 \quad \Leftrightarrow \quad \text{Cov}(\text{PC}_i, u_i) = 0 \tag{2}$$

¿Por qué PC está correlacionada con u?

La variable PC es **endógena** porque está correlacionada con variables no observadas en u_i:

$$\text{Cov}(\text{PC}_i, u_i) \neq 0 \tag{3}$$

Fuentes de endogeneidad:

1. **Habilidad no observada** ($\text{ability}_i \in u_i$):

$$\begin{aligned} \text{PC}_i &= f(\text{ability}_i, \ldots) \\ \text{GPA}_i &= g(\text{ability}_i, \ldots) \end{aligned} \tag{4}$$

   Estudiantes más hábiles $\Rightarrow$ mayor probabilidad de tener PC y mayor GPA.

2. **Antecedentes familiares** ($\text{family}_i \in u_i$):

$$\begin{aligned} \text{PC}_i &= f(\text{family}_i, \ldots) \\ \text{GPA}_i &= g(\text{family}_i, \ldots) \end{aligned} \tag{5}$$

   Mejor ambiente familiar $\Rightarrow$ mayor probabilidad de tener PC y mayor GPA.

3. **Motivación no observada** ($\text{motivation}_i \in u_i$):

$$\begin{aligned} \text{PC}_i &= f(\text{motivation}_i, \ldots) \\ \text{GPA}_i &= g(\text{motivation}_i, \ldots) \end{aligned} \tag{6}$$

   Mayor motivación $\Rightarrow$ buscar herramientas (PC) y mejor rendimiento (GPA).

Consecuencia:

$$\text{plim}\, \hat{\beta}_1^{\text{MCO}} = \beta_1 + \frac{\text{Cov}(\text{PC}, u)}{\text{Var}(\text{PC})} \neq \beta_1 \tag{7}$$

El estimador de MCO es **inconsistente** (sesgado incluso con $n \to \infty$).

1.2. Ejercicio 1b: Ingreso parental como instrumento

Condiciones para una Variable Instrumental válida z:

1. **Relevancia (testeable):**

$$\text{Cov}(z_i, \text{PC}_i) \neq 0 \tag{8}$$

   El instrumento debe estar correlacionado con la variable endógena.

2. **Exogeneidad/Restricción de exclusión (NO testeable):**

$$\text{Cov}(z_i, u_i) = 0 \tag{9}$$

   El instrumento NO debe estar correlacionado con el error, excepto a través de su efecto sobre x.

Evaluación del ingreso parental como IV para PC:

Sea $z_i = \text{ingreso\_parental}_i$

1. ¿Cumple RELEVANCIA?

$$\text{Cov}(\text{ingreso\_parental}_i, \text{PC}_i) > 0 \quad \checkmark \tag{10}$$

SÍ cumple: Familias con mayor ingreso tienen mayor probabilidad de comprar PCs. Podemos verificarlo con la regresión de primera etapa:

$$\text{PC}_i = \pi_0 + \pi_1(\text{ingreso\_parental}_i) + v_i \tag{11}$$

Esperamos que $\pi_1 > 0$ y estadísticamente significativo.

2. ¿Cumple EXOGENEIDAD?

$$\text{Cov}(\text{ingreso\_parental}_i, u_i) \stackrel{?}{=} 0 \tag{12}$$

NO cumple: El ingreso parental afecta al GPA a través de múltiples canales:

$$\text{GPA}_i = \beta_0 + \beta_1\text{PC}_i + \gamma_1\text{tutorías}_i + \gamma_2\text{libros}_i + \gamma_3\text{ambiente}_i + \tilde{u}_i \tag{13}$$

donde todas las variables tutorías_i, libros_i, ambiente_i están en u_i y son afectadas por el ingreso parental:

$$\begin{aligned} \text{tutorías}_i &= f(\text{ingreso\_parental}_i) \\ \text{libros}_i &= f(\text{ingreso\_parental}_i) \\ \text{ambiente}_i &= f(\text{ingreso\_parental}_i) \end{aligned} \tag{14}$$

Por lo tanto:

$$\text{Cov}(\text{ingreso\_parental}_i, u_i) \neq 0 \quad \times \tag{15}$$

Conclusión:

El ingreso parental **NO** es un instrumento válido para PC porque:

✓ Cumple relevancia

× **NO** cumple exogeneidad

1.3. Ejercicio 1c: Grants aleatorios como Variable Instrumental

Aleatorización como fuente de exogeneidad:

Si una variable z se asigna de manera completamente aleatoria, entonces por construcción:

$$z_i \perp\!\!\!\perp u_i \quad \Rightarrow \quad \text{Cov}(z_i, u_i) = 0 \tag{16}$$

La aleatorización garantiza que z es independiente de todas las características observadas y no observadas.

Construcción de la variable instrumental:

Sea GRANT_i una variable binaria:

$$\text{GRANT}_i = \begin{cases} 1 & \text{si el estudiante } i \text{ recibió el apoyo hace 4 años} \\ 0 & \text{si el estudiante } i \text{ NO recibió el apoyo} \end{cases} \tag{17}$$

¿Por qué GRANT es un buen instrumento?

1. EXOGENEIDAD (por aleatorización):

Como los grants se asignaron **al azar**:

$$\text{GRANT}_i \perp\!\!\!\perp (u_i, \text{ability}_i, \text{family}_i, \ldots) \quad \Rightarrow \quad \text{Cov}(\text{GRANT}_i, u_i) = 0 \quad \checkmark \tag{18}$$

2. RELEVANCIA:

Los estudiantes que recibieron el grant tienen mayor probabilidad de tener PC:

$$\text{Cov}(\text{GRANT}_i, \text{PC}_i) > 0 \quad \checkmark \tag{19}$$

Podemos verificarlo con la regresión de primera etapa:

$$\text{PC}_i = \pi_0 + \pi_1 \text{GRANT}_i + v_i \tag{20}$$

Esperamos que $\pi_1 > 0$ y estadísticamente significativo.

Estrategia de estimación:

Opción 1: Estimador de Wald (para instrumento binario)

Dado que GRANT es binario, podemos usar:

$$\hat{\beta}_1^{\text{Wald}} = \frac{\overline{\text{GPA}}_1 - \overline{\text{GPA}}_0}{\overline{\text{PC}}_1 - \overline{\text{PC}}_0} \tag{21}$$

donde:

- $\overline{\text{GPA}}_1$ = GPA promedio del grupo con GRANT = 1
- $\overline{\text{GPA}}_0$ = GPA promedio del grupo con GRANT = 0
- $\overline{\text{PC}}_1$ = proporción con PC en el grupo con GRANT = 1
- $\overline{\text{PC}}_0$ = proporción con PC en el grupo con GRANT = 0

Opción 2: Two-Stage Least Squares (2SLS)

Primera etapa:

$$\text{PC}_i = \pi_0 + \pi_1 \text{GRANT}_i + v_i \tag{22}$$

Obtenemos los valores ajustados:

$$\widehat{\text{PC}}_i = \hat{\pi}_0 + \hat{\pi}_1 \text{GRANT}_i \tag{23}$$

Segunda etapa:

$$\text{GPA}_i = \beta_0 + \beta_1 \widehat{\text{PC}}_i + \varepsilon_i \tag{24}$$

El coeficiente $\hat{\beta}_1$ de la segunda etapa es el estimador IV.

Interpretación:

$\hat{\beta}_1$ estima el **Local Average Treatment Effect (LATE)**:

$$\text{LATE} = E[\text{GPA}_i(1) - \text{GPA}_i(0) \mid \text{complier}_i] \tag{25}$$

donde los “compliers” son los estudiantes cuya decisión de tener PC fue afectada por recibir el grant.

2. EJERCICIO 2: Derivaciones del Estimador IV

2.1. Ejercicio 2a: Fórmula general del estimador IV

Modelo y supuestos:

Queremos estimar:

$$y_i = \beta_0 + \beta_1 x_i + u_i \tag{26}$$

Problema: $\text{Cov}(x_i, u_i) \neq 0$ (endogeneidad)

Tenemos un instrumento z_i que satisface:

- **Exogeneidad:** $\text{Cov}(z_i, u_i) = 0$
- **Relevancia:** $\text{Cov}(z_i, x_i) \neq 0$

DERIVACIÓN DEL ESTIMADOR IV:

Paso 1: Partir de los momentos poblacionales

Multiplicamos el modelo por z_i y tomamos esperanza:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + u_i \\ z_i \cdot y_i &= z_i \cdot \beta_0 + z_i \cdot \beta_1 x_i + z_i \cdot u_i \\ E[z_i y_i] &= E[z_i]\beta_0 + \beta_1 E[z_i x_i] + E[z_i u_i] \end{aligned} \tag{27}$$

Paso 2: Usar el supuesto de exogeneidad

Como $\text{Cov}(z_i, u_i) = 0$:

$$E[z_i u_i] = E[z_i]E[u_i] = 0 \quad \text{(porque } E[u_i] = 0) \tag{28}$$

Por lo tanto:

$$E[z_i y_i] = E[z_i]\beta_0 + \beta_1 E[z_i x_i] \tag{29}$$

Restando $E[z_i]E[y_i]$ a ambos lados:

$$\begin{aligned} E[z_i y_i] - E[z_i]E[y_i] &= E[z_i]\beta_0 + \beta_1 E[z_i x_i] - E[z_i][\beta_0 + \beta_1 E[x_i]] \\ \text{Cov}(z_i, y_i) &= \beta_1 [E[z_i x_i] - E[z_i]E[x_i]] \\ \text{Cov}(z_i, y_i) &= \beta_1 \text{Cov}(z_i, x_i) \end{aligned} \tag{30}$$

Paso 3: Despejar β_1

$$\boxed{\beta_1 = \frac{\text{Cov}(z_i, y_i)}{\text{Cov}(z_i, x_i)}} \tag{31}$$

Paso 4: Análogo muestral (estimador)

Usando el método de momentos, reemplazamos las covarianzas poblacionales por sus análogos muestrales:

$$\boxed{\hat{\beta}_1^{\text{IV}} = \frac{\widehat{\text{Cov}}(z_i, y_i)}{\widehat{\text{Cov}}(z_i, x_i)} = \frac{\sum_{i=1}^n (z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})}} \tag{32}$$

Propiedades del estimador:

1. **Consistencia:**
$$\text{plim}_{n\to\infty} \hat{\beta}_1^{\text{IV}} = \beta_1 \tag{33}$$

2. **Insesgamiento asintótico:**
$$\lim_{n\to\infty} E[\hat{\beta}_1^{\text{IV}}] = \beta_1 \tag{34}$$

3. **Distribución asintótica:**
$$\sqrt{n}(\hat{\beta}_1^{\text{IV}} - \beta_1) \xrightarrow{d} N(0, V) \tag{35}$$

   donde V depende de la correlación entre z y x.

2.2. Ejercicio 2b: Estimador de Wald para instrumento binario

Caso especial: Instrumento binario

Cuando el instrumento z_i es una variable binaria ($z_i \in \{0, 1\}$), el estimador IV se simplifica considerablemente.

OBJETIVO: Mostrar que cuando z_i es binario:

$$\hat{\beta}_1^{\text{IV}} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0} \tag{36}$$

donde:

- $\bar{y}_1 = \frac{1}{n_1} \sum_{i:z_i=1} y_i$ (promedio de y para $z = 1$)
- $\bar{y}_0 = \frac{1}{n_0} \sum_{i:z_i=0} y_i$ (promedio de y para $z = 0$)
- $\bar{x}_1, \bar{x}_0$ definidos similarmente
- $n_1 = \sum_i \mathbb{1}(z_i = 1)$, $n_0 = \sum_i \mathbb{1}(z_i = 0)$

DERIVACIÓN:
Paso 1: Calcular $E[z_i]$ cuando z es binaria
Sea $p = P(z_i = 1)$. Entonces:

$$E[z_i] = 0 \cdot P(z_i = 0) + 1 \cdot P(z_i = 1) = p \tag{37}$$

En la muestra:

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{n_1}{n} \tag{38}$$

Paso 2: Calcular Cov(z_i, y_i)

$$\begin{aligned} \text{Cov}(z_i, y_i) &= E[z_i y_i] - E[z_i]E[y_i] \\ &= E[y_i \mid z_i = 1] \cdot P(z_i = 1) - p \cdot E[y_i] \end{aligned} \tag{39}$$

Expandiendo $E[y_i]$:

$$E[y_i] = E[y_i \mid z_i = 1]P(z_i = 1) + E[y_i \mid z_i = 0]P(z_i = 0) \tag{40}$$

Sustituyendo:

$$\begin{aligned} \text{Cov}(z_i, y_i) &= p \cdot E[y_i \mid z_i = 1] - p[p \cdot E[y_i \mid z_i = 1] + (1-p) \cdot E[y_i \mid z_i = 0]] \\ &= p \cdot E[y_i \mid z_i = 1] - p^2 \cdot E[y_i \mid z_i = 1] - p(1-p) \cdot E[y_i \mid z_i = 0] \\ &= p(1-p) \cdot E[y_i \mid z_i = 1] - p(1-p) \cdot E[y_i \mid z_i = 0] \\ &= p(1-p)[E[y_i \mid z_i = 1] - E[y_i \mid z_i = 0]] \end{aligned} \tag{41}$$

En la muestra, con $p = n_1/n$:

$$\widehat{\text{Cov}}(z_i, y_i) = \frac{n_1}{n}\left(1 - \frac{n_1}{n}\right)(\bar{y}_1 - \bar{y}_0) = \frac{n_1 n_0}{n^2}(\bar{y}_1 - \bar{y}_0) \tag{42}$$

Paso 3: Calcular Cov(z_i, x_i) **de manera similar**
Por el mismo razonamiento:

$$\widehat{\text{Cov}}(z_i, x_i) = \frac{n_1 n_0}{n^2}(\bar{x}_1 - \bar{x}_0) \tag{43}$$

Paso 4: Sustituir en la fórmula del estimador IV

$$\begin{aligned} \hat{\beta}_1^{\text{IV}} &= \frac{\widehat{\text{Cov}}(z_i, y_i)}{\widehat{\text{Cov}}(z_i, x_i)} \\ &= \frac{\frac{n_1 n_0}{n^2}(\bar{y}_1 - \bar{y}_0)}{\frac{n_1 n_0}{n^2}(\bar{x}_1 - \bar{x}_0)} \\ &= \boxed{\frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0}} \end{aligned} \tag{44}$$

Esto es el **estimador de Wald**.

2.3. Interpretación del estimador de Wald

FORMA INTUITIVA:

$$\hat{\beta}_1^{\text{Wald}} = \frac{\text{Diferencia en } y \text{ entre grupos}}{\text{Diferencia en } x \text{ entre grupos}} = \frac{\text{Forma reducida}}{\text{Primera etapa}} \tag{45}$$

Componentes:

1. **Forma reducida (Reduced Form):**

$$\bar{y}_1 - \bar{y}_0 = \text{Efecto total del instrumento sobre el resultado} \tag{46}$$

2. **Primera etapa (First Stage):**

$$\bar{x}_1 - \bar{x}_0 = \text{Efecto del instrumento sobre el tratamiento} \tag{47}$$

Conexión con 2SLS:
El estimador de Wald es equivalente a Two-Stage Least Squares:
Etapa 1:

$$x_i = \pi_0 + \pi_1 z_i + v_i \tag{48}$$

Cuando z es binario:

$$\hat{\pi}_1 = \bar{x}_1 - \bar{x}_0 \tag{49}$$

Los valores ajustados son:

$$\hat{x}_i = \begin{cases} \bar{x}_1 & \text{si } z_i = 1 \\ \bar{x}_0 & \text{si } z_i = 0 \end{cases} \tag{50}$$

Etapa 2:

$$y_i = \beta_0 + \beta_1 \hat{x}_i + \varepsilon_i \tag{51}$$

El coeficiente de esta regresión es exactamente el estimador de Wald.

2.4. Aplicación al Ejercicio 1c

Aplicación a los grants aleatorios:
En el contexto del Ejercicio 1c, donde:

- $y_i = \text{GPA}_i$
- $x_i = \text{PC}_i$
- $z_i = \text{GRANT}_i$

El estimador de Wald se convierte en:

$$\hat{\beta}_1^{\text{Wald}} = \frac{\overline{\text{GPA}}_{\text{grant=1}} - \overline{\text{GPA}}_{\text{grant=0}}}{\overline{\text{PC}}_{\text{grant=1}} - \overline{\text{PC}}_{\text{grant=0}}} \tag{52}$$

Interpretación:

- **Numerador:** Diferencia en GPA promedio entre quienes recibieron el grant y quienes no.
- **Denominador:** Diferencia en la proporción con PC entre quienes recibieron el grant y quienes no.
- **Cociente:** Efecto causal de tener una PC sobre el GPA (para los compliers).

Ejemplo numérico:
Supongamos que encontramos:

- $\overline{\text{GPA}}_{\text{grant}=1} = 3{,}2$
- $\overline{\text{GPA}}_{\text{grant}=0} = 3{,}0$
- $\overline{\text{PC}}_{\text{grant}=1} = 0{,}8$ (80 % tienen PC)
- $\overline{\text{PC}}_{\text{grant}=0} = 0{,}4$ (40 % tienen PC)

Entonces:

$$\hat{\beta}_1^{\text{Wald}} = \frac{3{,}2 - 3{,}0}{0{,}8 - 0{,}4} = \frac{0{,}2}{0{,}4} = 0{,}5 \tag{53}$$

Interpretación: Tener una PC aumenta el GPA en 0.5 puntos, en promedio, para los estudiantes cuya decisión de tener PC fue afectada por recibir el grant.

3. Resumen de conceptos clave

CONCEPTOS FUNDAMENTALES:

1. **Endogeneidad:** $\text{Cov}(x, u) \neq 0 \Rightarrow$ MCO es inconsistente
2. **Variable Instrumental válida:**
   - Relevancia: $\text{Cov}(z, x) \neq 0$
   - Exogeneidad: $\text{Cov}(z, u) = 0$
3. **Estimador IV general:**
$$\hat{\beta}_1^{\text{IV}} = \frac{\text{Cov}(z, y)}{\text{Cov}(z, x)}$$
4. **Estimador de Wald (z binario):**
$$\hat{\beta}_1^{\text{Wald}} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0}$$

5. **2SLS:**
   - Etapa 1: $\hat{x} = \hat{\pi}_0 + \hat{\pi}_1 z$
   - Etapa 2: $y = \hat{\beta}_0 + \hat{\beta}_1 \hat{x} + \varepsilon$
6. **LATE:** IV identifica el efecto para “compliers”