

## **Trabajo Práctico N° 5:** **Variables Instrumentales.**

### **Ejercicio 1.**

Suponer un modelo de regresión simple:

$$y_i = \beta_0 + \beta_1 x_i + u_i,$$

donde  $x_i$  es, potencialmente, endógena. Además, suponer que el instrumento,  $z_i$ , es una variable binaria. Mostrar que el estimador IV, en este caso, es:

$$\beta_1^{IV} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0},$$

donde  $\bar{y}_1$ ,  $\bar{x}_1$  ( $\bar{y}_0$ ,  $\bar{x}_0$ ) representan las medias cuando  $z=1$  ( $z=0$ ).

$$\beta^{IV} = (Z'X)^{-1} Z'y$$

$$\beta^{IV} = \begin{bmatrix} 1_n \\ z \end{bmatrix} \begin{bmatrix} 1_n & x \end{bmatrix}^{-1} \begin{bmatrix} 1_n \\ z \end{bmatrix} y$$

$$\beta^{IV} = \begin{bmatrix} 1_n' 1_n & 1_n' x \\ z' 1_n & z' x \end{bmatrix}^{-1} \begin{bmatrix} 1_n' y \\ z' y \end{bmatrix}$$

$$\beta^{IV} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n z_i & \sum_{i=1}^n z_i x_i \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n z_i y_i \end{bmatrix}$$

$$\beta^{IV} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ n_1 & \sum_{z_i=1} x_i \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{z_i=1} y_i \end{bmatrix}$$

$$\beta^{IV} = \frac{1}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} \begin{bmatrix} \sum_{z_i=1} x_i & -\sum_{i=1}^n x_i \\ -n_1 & n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{z_i=1} y_i \end{bmatrix}$$

$$\beta^{IV} = \begin{bmatrix} \frac{\sum_{z_i=1} x_i}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} & \frac{-\sum_{i=1}^n x_i}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} \\ \frac{-n_1}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} & \frac{n}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{z_i=1} y_i \end{bmatrix}.$$

$$\beta_1^{IV} = \frac{-n_1}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} \sum_{i=1}^n y_i + \frac{n}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i} \sum_{z_i=1} y_i$$

$$\beta_1^{IV} = \frac{n \sum_{z_i=1} y_i - n_1 \sum_{i=1}^n y_i}{n \sum_{z_i=1} x_i - n_1 \sum_{i=1}^n x_i}$$

$$\beta_1^{IV} = \frac{n \sum_{z_i=1} y_i - n_1 (\sum_{z_i=1} y_i + \sum_{z_i=0} y_i)}{n \sum_{z_i=1} x_i - n_1 (\sum_{z_i=1} x_i + \sum_{z_i=0} x_i)}$$

$$\beta_1^{IV} = \frac{n \sum_{z_i=1} y_i - n_1 \sum_{z_i=1} y_i - n_1 \sum_{z_i=0} y_i}{n \sum_{z_i=1} x_i - n_1 \sum_{z_i=1} x_i - n_1 \sum_{z_i=0} x_i}$$

$$\beta_1^{IV} = \frac{(n - n_1) \sum_{z_i=1} y_i - n_1 \sum_{z_i=0} y_i}{(n - n_1) \sum_{z_i=1} x_i - n_1 \sum_{z_i=0} x_i}$$

$$\beta_1^{IV} = \frac{n_0 \sum_{z_i=1} y_i - n_1 \sum_{z_i=0} y_i}{n_0 \sum_{z_i=1} x_i - n_1 \sum_{z_i=0} x_i}$$

$$\beta_1^{IV} = \frac{n_0 n_1 \bar{y}_1 - n_1 n_0 \bar{y}_0}{n_0 n_1 \bar{x}_1 - n_1 n_0 \bar{x}_0}$$

$$\beta_1^{IV} = \frac{n_0 n_1 (\bar{y}_1 - \bar{y}_0)}{n_0 n_1 (\bar{x}_1 - \bar{x}_0)}$$

$$\beta_1^{IV} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0}.$$

**Ejercicio 2.**

*En este ejercicio, se propone extender la simulación del Problem Set 1 a un marco en el que la asignación del tratamiento y quienes resultan tratados no son iguales.*

**(a)** *Inicializar una muestra con 100 observaciones. Generar resultados potenciales de no recibir el tratamiento como:  $Y_0 \sim \mathcal{N}(100, 30)$ .*

**Stata.**

**(b)** *Generar, ahora, un efecto de tratamiento constante e igual a 20 para todos, es decir,  $TE_i = 20$ , para todo  $i = 1, \dots, n$ . Generar una variable aleatoria normal estándar. Generar una variable de tratamiento  $D_i$  igual a 1 para aquellas observaciones que poseen un valor positivo en la variable aleatoria normal.*

**Stata.**

**(c)** *Generar una aleatoria uniforme en el intervalo  $[0, 1]$ . Con ella, generar variables que indiquen el tipo de individuo. Utilizar: *always taker* si la variable es menor a 0,25, *never taker* si la variable está entre 0,25 y 0,5, *defier* si la variable está entre 0,5 y 0,75 y *complier* si la variable es mayor a 0,75. Generar la variable de si los individuos toman el tratamiento o no dependiendo del grupo en el que están.*

**Stata.**

**(d)** *Generar la variable  $Y$  observada como  $Y = DY_1 + (1 - D) Y_0$ .*

**Stata.**

**(e)** *Estimar el LATE y comparar con el ATE.*

**Stata.**

### **Ejercicio 3.**

Leer el artículo “Property Rights for the Poor: Effects of Land Titling” de Galiani & Schargrotsky.

**(a)** ¿Qué efectos intentan estimar en el paper?

**(b)** ¿Cuál es la estrategia de identificación? ¿Por qué no funciona la diferencia de medias simple?

**(c)** Replicar los resultados del paper.

**Stata.**