

ECONOMETRÍA AVANZADA - TRABAJO PRÁCTICO N°3

Variables instrumentales

Fecha de entrega: 3 de mayo

Profesor: Walter Sosa Escudero
Asistente: Gastón García Zavaleta

Modalidad: enviar el informe en formato pdf y el script de R por mail a Gastón García Zavaleta (gzavaleta@udesa.edu.ar). **Los nombres de ambos archivos y el asunto del mail** deben ser “TP3 Econometría Avanzada - APELLIDOS”. Esperar mail confirmando la recepción.

Reglas de formato y presentación

- El objetivo de este trabajo práctico es revisar algunas dificultades empíricas en la estimación de modelos ante la presencia de endogeneidad, además de repasar conceptos teóricos básicos sobre variables instrumentales.
 - Incluyendo tablas, no debe exceder las ocho carillas A4. Se espera una redacción cuidada y profesional, prestando atención a aspectos estéticos en el diseño de tablas y en la presentación elegante de los resultados, como si se tratase de un verdadero trabajo académico o de consultoría profesional.
 - El trabajo debe elaborarse en los grupos conformados.
-

Ejercicio 1 (1 punto)

Consideren los siguientes modelos:

$$y_1 = z_1\delta_1 + x_1\delta_2 + u_1 \tag{1}$$

$$x_1 = z\Pi + u_2 \tag{2}$$

donde $z_1 \subset z$.

¿Cómo estimarían δ_2 usando 2SLS? ¿Cuál es el instrumento?

Ejercicio 2 (2 puntos)

Supongamos que hace cuatro años les proveyeron una base de datos $\{(y_i, x_i, z_i, u_i)\}_{i=1}^n$ generada por una muestra aleatoria de la distribución conjunta de (y, x, z, u) . La relación entre las variables está dada por el modelo estructural:

$$y = \beta_0 + \beta_1 x + u$$

En su momento, les dijeron que las siguientes cosas eran ciertas:

- No pueden suponer que $E(u|x) = 0$
- $Cov(u, z) = 0$
- $Cov(x, z) \neq 0$
- Hay variabilidad en los datos tanto para x como para z .

Como desconocían la metodología de variables instrumentales, ustedes corrieron “todas las regresiones por MCO posibles”, estimando una regresión de y en x , una de y en z y finalmente una de x en z . Los coeficientes de esas regresiones fueron:

$$\hat{\omega}_{yx} = \frac{\widehat{Cov(y, x)}}{\widehat{V(x)}}$$

$$\hat{\omega}_{yz} = \frac{\widehat{Cov(y, z)}}{\widehat{V(z)}}$$

$$\hat{\omega}_{xz} = \frac{\widehat{Cov(x, z)}}{\widehat{V(z)}}$$

Hoy en día, ustedes solo cuentan con los resultados de esas estimaciones dado que eliminaron la base de datos.

- a. ¿Pueden construir un estimador consistente para β_1 con los datos que tienen? Si la respuesta es sí, muestren que es consistente.
- b. Si les facilitaran la base de datos, ¿cómo sería el estimador de VI que usarían? Prueben que es consistente.

Ejercicio 3 (7 puntos)

Este ejercicio está basado en el paper de Angrist y Kruger de 1991 sobre retornos a la educación. Ellos emplean el trimestre de nacimiento como instrumento para la educación con el objetivo de estimar el impacto de la obligatoriedad escolar en los ingresos. La base de datos, QOB.dta, contiene las siguientes variables:

Nombre	Descripción
age	Edad
ageq	Edad medida en trimestres
ageqsq	Edad medida en trimestres, al cuadrado
educ	Número de años completos de educación
lwage	Logaritmo del salario mensual
married	Dummy de matrimonio (=1 si casado)
race	Dummy de raza (=1 si negro)
smsa	Dummy de ciudad central (=1 si reside en una ciudad central)
qob	Trimestre de nacimiento
yob	Año de nacimiento
census	Fuente de los datos

- Presenten y comenten un cuadro con estadísticas descriptivas de las variables que utilicen (no presenten estadísticos irrelevantes o que no planeen comentar). En particular, calculen y comenten el número promedio de años de escolaridad completa para cada “*quarter of birth*” para los hombres nacidos en las décadas del 30 y 40. ¿Qué observan?
- De ahora en adelante, usen únicamente datos de hombres nacidos entre 1930 y 1939 inclusive. Estimen los retornos a la educación mediante MCO: regresen el logaritmo de los salarios mensual contra la educación, la edad en trimestres, la edad en trimestres al cuadrado, la raza y las dummies de estado civil y ciudad central. Interpreten los resultados con exactitud (es decir, cada uno de los parámetros estimados) y expliquen por qué los retornos a la educación estimados pueden llegar a no tener una interpretación causal.
- Expliquen por qué “*quarter of birth*” podría ser una buena variable instrumental para la educación en la ecuación de salarios.
- Supongan que tenemos tres variables instrumentales: Z_1 , Z_2 y Z_3 , representando dummies para los que nacieron en el primer, segundo y tercer trimestre del año respectivamente. Generen estos instrumentos.
- Testeen la hipótesis de que las dummies trimestrales no tienen efecto conjunto sobre los años de educación.
- Teniendo en cuenta que buscamos estimar los retornos a la educación utilizando los instrumentos del punto (d) y como controles las variables **ageq**, **ageqsq**, **race**, **married** y **smsa**, describan y estimen la ecuación “*first-stage*”. Comenten los resultados de la regresión.
- Estudien la posibilidad de instrumentos débiles en el modelo sobre identificado.
- Estimen los retornos a la educación por Mínimos Cuadrados en 2 Etapas (MC2E) y comparen los resultados con los obtenidos en el punto (b).

Referencias

Angrist, J. and A.B. Krueger (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *Quarterly Journal of Econometrics*, Vol. 106, No. 4, pp. 979 1014.