

Teoría Econométrica

Problem Set 4

Profesor: Tatiana Rosá

Ayudante: Alejo Eyzaguirre

Noviembre 2021

Fecha de entrega: Miércoles 24 de noviembre 17:00 por buzón en canvas.

Instrucciones: Se debe realizar uno de los ejercicios propuestos a elección.

1 Retornos a la educación

Sea y_i el logaritmo de salario (por hora) de un individuo i , y sea x_i su nivel educativo (en años). Queremos saber el efecto *causal* de x_i en y_i .

Empezamos por asumir el siguiente modelo:

$$y_i = \delta + \alpha x_i + u_i$$

$$x_i = \mu + \beta z_i + v_i$$

$$u_i = z_i + \eta_i$$

En este sistema z_i es inobservable, iid y con varianza σ_z^2 . Los errores v_i y η_i son iid, incorrelacionados, tienen media cero y varianza σ_v^2 y σ_η^2 respectivamente. Finalmente, $\mathbb{E}(v_i z_i) = \mathbb{E}(\eta_i z_i) = 0$

Se le pide:

1. Interprete el modelo estructural ¿Qué interpretación le da a z_i ?
2. Sea $\hat{\alpha}$ el coeficiente de la proyección lineal de y_i sobre x_i . Compute:

$$\text{plim } \hat{\alpha} - \alpha$$

Ayuda: Piense en $\hat{\alpha}$ como $\hat{\alpha} = \frac{\text{Cov}(x_i y_i)}{\text{Var}(x_i)}$ ¿Qué signo espera para esta diferencia? Discuta.

3. De la lista de variables que se presenta a continuación discuta cuales son potenciales instrumentos válidos y porqué.
 - Coeficiente Intelectual
 - Distancia a la Universidad
 - Profesión del padre
 - Salario mensual
 - Horas trabajadas
 - Mes de nacimiento
4. Utilizando un solo instrumento, explique como estimar consistentemente α . Demuestre la consistencia del estimador propuesto.

2 Endogeneidad y error de medida:

Un investigador que también está intentando estimar el efecto causal de la educación sobre los salarios, utilizando un instrumento válido encuentra:

$$\tilde{\alpha} = 1.2\hat{\alpha}$$

Para intentar explicar este resultado, consideramos este segundo modelo aumentado:

$$y_i = \delta + \alpha x_i + u_i$$

$$x_i = \mu + \beta z_i + v_i$$

$$u_i = z_i + \eta_i$$

$$\tilde{y}_i = y_i + \epsilon_i$$

$$\tilde{x}_i = x_i + \nu_i$$

donde \tilde{y}_i y \tilde{x}_i son las **únicas variables observadas del modelo**.

Además de los supuestos del ejercicio anterior, asumimos que ϵ_i y ν_i son iid, media cero, incorrelacionados y que $\mathbb{E}(\epsilon_i z_i) = \mathbb{E}(\epsilon_i v_i) = \mathbb{E}(\nu_i z_i) = \mathbb{E}(\nu_i v_i) = \mathbb{E}(\nu_i \eta_i) = 0$. Asumimos además que ν_i tiene varianza σ_ν^2 .

Se le pide:

1. Interprete este nuevo modelo. ¿Cuál es la diferencia con el modelo presentado en el ejercicio 1?
2. Sea $\hat{\alpha}$ el coeficiente de la proyección lineal de \tilde{y}_i sobre \tilde{x}_i . Compute:

$$\text{plim } \hat{\alpha} - \alpha$$

3. En este nuevo modelo, ¿estimar por dos MC2E, nos permite estimar consistentemente α ? Comente que condiciones adicionales deberían cumplirse (si alguna) respecto al caso del ejercicio 1 (donde tenemos solamente endogeneidad por omisión de variable relevante).
4. Se cree que en los datos utilizados por el investigador para esta estimación el error de medida de los años de educación representa el 10% de la varianza de esta variable. ¿Cómo relaciona ud. esto con el sesgo encontrado por el investigador en el estimador por MCO de α ($\hat{\alpha}$). Discuta.

3 Instrumentos débiles

Considere el siguiente modelo:

$$y_i = \beta x_i + u_i$$

donde β y x_i son escalares.

Sea z_i un escalar tal que: $x_i = \gamma z_i + v_i$

Se asume además que los datos son iid, que la $\text{Var}(z_i) \neq 0$, que $\mathbb{E}(v_i | z_i) = 0$ y que $\mathbb{E}(u_i | z_i) = 0$.

Sea $\hat{\beta}_{2SLS}$ el estimador en dos etapas de y_i sobre x_i usando z_i como instrumento.

Instrumento fuerte:

Comenzamos asumiendo que $\gamma \neq 0$ es una constante.

Se le pide:

1. Pruebe que $\mathbb{E}(x_i z_i) \neq 0$
2. Pruebe que $\hat{\beta}$ converge en probabilidad a β cuando N tiende a infinito

Instrumento débil:

Asumimos ahora que $\gamma = \frac{1}{\sqrt{N}}$

Se le pide:

1. Pruebe que

$$plim \frac{1}{N} \sum_{i=1}^N z_i x_i = 0$$

2. Pruebe que

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N z_i x_i \xrightarrow{d} N(\mu, V)$$

donde debe especificar μ y V

3. Compute $AVar(\hat{\beta}_{2SLS})$. ¿Qué pasa cuando $N \rightarrow \infty$? Comente su implicación sobre la posibilidad de hacer inferencia usando esta estimación.
4. Relajamos levemente supuesto de exogeneidad de z_i tal que $\mathbb{E}(u_i|z_i) = 0.001$. ¿Qué pasa con la consistencia del estimador? ¿Pasaría lo mismo si fuera un instrumento fuerte? Comente.