

ECONOMETRÍA - TRABAJO PRÁCTICO N°3

---

# Variables instrumentales: Consumo de cigarrillos y peso al nacer

---

Profesor: Walter Sosa Escudero  
Asistente: Florencia Hnilo

**Modalidad:** enviar el informe en formato pdf por mail a Florencia Hnilo ([fhnilo@udesa.edu.ar](mailto:fhnilo@udesa.edu.ar)) con el asunto "TP3 Econometría Exactas - APELLIDOS". Además, adjuntar el script de R con el código utilizado. Esperar mail confirmando la recepción.

## Reglas de formato y presentación

- El objetivo del presente trabajo práctico es aplicar el método de estimación con variables instrumentales.
- Incluyendo tablas, no debe exceder las cinco 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 grupos de dos o tres personas, que entregarán una sola copia por grupos.

---

Hasta hace unos años se pensaba que el ambiente intrauterino era capaz de proteger a los humanos en gestación de factores perjudiciales del entorno. Sin embargo, investigaciones más recientes señalan lo contrario: los factores ambientales pueden incidir fuertemente en la salud del feto. En este trabajo vamos a analizar la relación entre peso al nacer, variable comúnmente usada como medida de la salud durante la gestación, y el consumo de cigarrillo por parte de la madre durante ese período<sup>1</sup>. Para ello utilizaremos la base de datos `tp3_iv.csv` subida a la página del curso, la cual contiene datos sobre 1388 nacidos en Estados Unidos en 1988 y sus madres. Las variables incluidas son:

---

<sup>1</sup>Currie, Neidell y Schmieder (2008) y Currie (2011) han tratado este tema.

Variable	Descripción
lfaminc	Logaritmo del ingreso familiar
cigtax	Impuesto al cigarrillo en el estado de residencia
cigprice	Precio del paquete de cigarrillo en el estado de residencia
lbwght	Logaritmo del peso al nacer
fatheduc	Años de educación del padre
motheduc	Años de educación de la madre
parity	Orden de nacimiento
male	1=hombre, 0=mujer
white	1=blanco, 0 en otro caso
packs	Paquetes de cigarrillos fumados por día durante el embarazo

1. Presente y comente un cuadro con estadística descriptiva de las variables que utilice (no presente estadísticos irrelevantes o que no planee comentar).
2. Plantee un modelo que tenga como variable dependiente el logaritmo del peso al nacer y como variable explicativa la cantidad de paquetes de cigarrillos consumidos diariamente durante el embarazo. Estime este modelo con el método de mínimos cuadrados ordinarios (MCO). Luego expanda este modelo incorporando como controles las variables sexo, orden de nacimiento, logaritmo del ingreso familiar. Estime este segundo modelo usando el mismo método. Comente los resultados.
3. Indique razones por las que la variable *packs* podría estar correlacionada con el término de error en el segundo modelo. ¿Cómo afectaría esto a las estimaciones obtenidas?
4. ¿Será el precio de los paquetes de cigarrillos (*cigprice*) un posible instrumento para la variable *packs*? ¿Qué razones podrían esgrimirse a favor y en contra?
5. ¿Por qué incluiría esta variable como instrumento y no como un control adicional en la regresión? Es decir, en términos generales, ¿cuándo una variable debe incluirse como control y cuándo como instrumento?
6. Estime el segundo modelo usando el método de mínimos cuadrados en dos etapas (MC2E) con *cigprice* como instrumento. Compare esta estimación con la hecha usando el método MCO.
7. Implemente el método MC2E paso a paso. ¿Qué regresores ha incluido en la primera etapa? ¿Sólo *cigprice* o también los controles del modelo a estimar? ¿Por qué? ¿Qué ocurriría si lo hiciera de otro modo? Tras estimar los parámetros de la segunda etapa, ¿obtuvo las mismas estimaciones de coeficientes y errores estándar que en el punto 6?
8. A la luz de lo desarrollado en el punto 7, ¿qué podría decir sobre la validez de lo hallado en el punto 6? ¿Qué valor toma el estadístico F de la regresión primera etapa? ¿Es este el estadístico relevante para estudiar instrumentos débiles? Considere la frase “Es comúnmente aceptado que si el estadístico F de la primera etapa es mayor a 10, podemos estar tranquilos al asumir que no estamos ante instrumentos débiles”. ¿Es correcta esta afirmación? ¿Cómo debe usarse la estimación de la primera etapa para analizar la posibilidad de que el instrumento sea débil?
9. Ahora considere el impuesto al cigarrillo (*cigtax*), ¿qué razones hay para pensar que

podría usarse esta variable como instrumento? Estime el segundo modelo del punto 2 empleando *cigprice* y *cigtax* como instrumentos, usando GMM (tanto con matriz de ponderadores robusta como no robusta a heterocedasticidad) y con el método MC2E. Interprete y compare los resultados. ¿Se obtienen las mismas estimaciones de los coeficientes y de los errores estándar?

10. Aplique el test de sobre-identificación para el caso de estos dos instrumentos, basado en la estimación de GMM robusta. ¿Qué resultados obtiene? ¿Qué conclusión puede sacar en base a los resultados de este test?
11. Estudie la posibilidad de instrumentos débiles en el modelo sobre identificado.

### Referencias

Currie, J. (2011). "Inequality at birth: Some causes and consequences". *American Economic Review: Papers Proceedings*, 101(3), 1-22. Disponible en <https://www.nber.org/papers/w16798>

Currie, J., Neidell, M. y Schmieder, J. (2008). "Air pollution and infant health: Lessons from New Jersey" NBER Working paper N° 14196. Disponible en <https://www.nber.org/papers/w14196>

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*. MIT Press.