

---

# **Ejemplo de estimación con variables instrumentales**

# Educación y salario de mujeres casadas

---

Utilizamos los datos sobre mujeres trabajadoras de un grupo de mujeres casadas. Se quiere estimar el rendimiento de los años de educación formal en el modelo de regresión simple.

- *educacion*: los años de educación formal;
- *lsalario*: logaritmo de la variable salario;
- *trabajadora*=1 si trabaja fuera de casa;
- *educpadre*: los años de educación del padre.

# Estimación por MCO

---

- El modelo de regresión simple

$$\log(\textit{salario}) = \beta_0 + \beta_1 \textit{educacion} + \epsilon$$

- Modelo estimado con MCO:

$$\log(\widehat{\textit{salario}}) = \begin{array}{cc} -0.185 & +0.109\textit{educacion} \\ (0.185) & (0.014) \end{array}$$

$$n = 428 \quad R^2 = 0.118$$

- La estimación de  $\beta_1$  implica un porcentaje de casi el 11% más de salario por cada año extra de educación formal.

# Estimación por VI

---

- Usamos la variable *educpadre* como variable instrumental para *educacion*. ¿Qué suponemos?

# Estimación por VI

---

- Usamos la variable *educpadre* como variable instrumental para *educacion*. ¿Qué suponemos?
  - *educpadre* está incorrelacionada con el error  $\epsilon$

# Estimación por VI

---

- Usamos la variable *educpadre* como variable instrumental para *educacion*. ¿Qué suponemos?
  - *educpadre* está incorrelacionada con el error  $\epsilon$
  - *educacion* y *educpadre* están correlacionadas y por lo tanto podemos hacer una regresión simple

# Estimación por VI

---

- Usamos la variable *educpadre* como variable instrumental para *educacion*. ¿Qué suponemos?
  - *educpadre* está incorrelacionada con el error  $\epsilon$
  - *educacion* y *educpadre* están correlacionadas y por lo tanto podemos hacer una regresión simple

$$\widehat{educacion} = 10.24 + 0.269educpadre$$

■  $(0.28) \quad (0.029)$

$$n = 428 \quad R^2 = 0.173$$

- El estadístico  $t$  para la variable *educpadre* es  $t = 9.28$  que indica que *educacion* y *educpadre* tienen una relación positiva estadísticamente significativa.

# Estimación por VI

---

Ahora usando  $\widehat{educacion}$  tenemos el modelo ajustado:

$$\log(\widehat{salario}) = 0.441 + 0.059\widehat{educacion}$$

■  $(0.446) \quad (0.035)$

$$n = 428 \quad R^2 = 0.093$$

- El salario se incrementa en un 5.9% por año adicional de educación cuando usamos esta VI.
- Esto sugiere que la estimación MCO es demasiado alta y es consistente con el sesgo por omisión de la variable *capacidad*.

# Conclusiones

---

- Pero deberíamos recordar que éstas son estimaciones de una única muestra: nunca podremos saber si 0.109 está por encima del verdadero efecto, o si 0.059 está más cerca del mismo.
- Además el error estándar de la estimación VI es dos veces y media mayor que el de la estimación MCO.
- El intervalo de confianza al 95% para  $\beta_1$  con MCO es más ajustado.
- También el intervalo de confianza al 95% del estimador VI contiene la estimación por MCO.
- Por lo tanto no podemos decir si las diferencias entre las dos estimaciones son estadísticamente significativas.