

Hoja de Ejercicios 3 Especificación y Problemas de los Datos

ECONOMETRÍA I. UC3M

1. [W 9.1] El R^2 de estimar el modelo

$$\log(\text{salary}) = \beta_0 + \beta_1(\text{sales}) + \beta_2 \log(\text{mktval}) + \beta_3 \text{profit margin} \\ + \beta_4 \text{ceoten} + \beta_5 \text{comten} + u$$

usando los datos CEOSAL2 es $R^2 = 0,353$ ($n = 177$). Si se añaden al modelo ceoten^2 y comten^2 entonces $R^2 = ,375$. ¿Hay evidencia de incorrecta especificación funcional en el modelo?

2. [W. 9.2] Para una muestra de institutos en el estado de Michigan, la variable math10 es el porcentaje de aprobados en un test estandarizado de matemáticas. Estamos interesados en estimar el efecto sobre los resultados en matemáticas del gasto por estudiante, para lo que se plantea el modelo

$$\text{math10} = \beta_0 + \beta_1 \log(\text{expend}) + \beta_2 \log(\text{enroll}) + \beta_3 \text{poverty} + u,$$

donde poverty es el porcentaje de estudiantes que viven bajo el umbral de pobreza, expend es el gasto por estudiante y enroll el número de estudiantes en el instituto.

- a) La variable lnchprg es el porcentaje de estudiantes que pueden solicitar el programa federal de ayudas para el comedor. ¿Por qué lnchprg es una variable proxy sensata para poverty ?
- b) Esta tabla proporciona los estimadores MCO, con y sin lnchprg como variable explicativa:

VARIABLES EXPLICATIVAS	(1)	(2)
$\log(\text{expend})$	11,13 (3,30)	7,75 (3,04)
$\log(\text{enroll})$	0,022 (,615)	-1,26 (,58)
lnchprg	-	-0,324 (,036)
intercept	-69,24 (26,74)	-23,14 (24,99)
N	428	428
R^2	0,0293	0,1893

Explique por qué el efecto de los gastos sobre math10 es menor en la columna (2) que en la columna (1). ¿Es este efecto en la columna (2) estadísticamente mayor que cero?

- c) ¿Es la tasa de aprobados mayor en los institutos grandes, ceteribus paribus? Justifique su respuesta.
- d) ¿Qué conclusión se obtiene del sustancial incremento del R^2 de la columna (1) a la (2)?

3. [W. 9.4] La siguiente ecuación caracteriza las horas semanales de televisión que ve un niño en función de su edad, educación de la madre, educación del padre y número de hermanos:

$$tvhours^* = \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 motheduc + \beta_4 fatheduc + \beta_5 sibs + u.$$

Cabe sospechar que $tvhours^*$ esté medida con error en nuestra encuesta. Si $tvhours$ es el número de horas de televisión declaradas por semana:

- a) ¿Qué requiere el supuesto clásico de errores en variables en esta aplicación?
- b) ¿Es verosímil que se cumpla el supuesto clásico de errores en variables en este caso? Justifique su respuesta.

4. [W. 9.6] Estime la educación

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 roe + \beta_3 rosneg + u$$

usando los datos CEOSAL1.

- a) Para estudiar la especificación, aplique un contraste RESET. ¿Hay evidencia de especificación errónea?
- b) Compute una versión del contraste RESET robusta a heteroscedasticidad. ¿Cambian sus conclusiones?
- c) Considere si este nuevo modelo es más apropiado que el anterior:

$$\log(salary) = \beta_0 + \beta_1 sales + \beta_2 sales^2 + \beta_3 roe + \beta_4 roe * rosneg + u,$$

efectuando los contrastes que considere adecuados.