EXAMEN EXTRAORDINARIO DE ECONOMETRÍA

Universidad Carlos III de Madrid

Junio 2016

Responda a las 4 preguntas en dos horas y media.

1. *Se está llevando a cabo una investigación econométrica sobre los precios de las casas vendidas recientemente en un área urbana. Los datos de muestra se componen de 88 observaciones sobre las siguientes variables.*

$$
\begin{aligned}
P_i &= \text{precio de venta de la casa } i\text{-ésima en miles de dólares.}\\
HS_i &= \text{tamaño de la casa } i\text{-ésima en cientos de pies cuadrados.}\\
YS_i &= \text{tamaño de la parcela de la casa } i\text{-ésima en cientos de pies cuadrados.}\\
DC_i &= \begin{cases} 1 \text{ si la casa } i\text{-ésima es de estilo colonial} \\ 0 \text{ en otro caso} \end{cases}
\end{aligned}
$$

El modelo lineal propuesto es

$$
\begin{aligned}
\ln P_i &= \beta_0 + \beta_1 \ln HS_i + \beta_2 \ln YS_i + \beta_3 (\ln HS_i)^2 + \beta_4 (\ln YS_i)^2 + \beta_5 (\ln HS_i)(\ln YS_i)\\
&\quad + \beta_6 DC_i + U_i,
\end{aligned}
$$

donde los β_j's son coeficientes de regresión y U_i es un término de error, que se supone satisface los supuestos clásicos. La siguiente tabla muestra la estimación del modelo bajo diferentes restricciones sobre los coeficientes (errores estándar entre paréntesis)

Explanatory Var.	(1)	(2)	(3)	(4)
Intercept	9.321 (2.500)	2.767 (0.593)	8.456 (2.133)	2.639 (0.244)
$\ln HS_i$	-3.029 (1.398)	0.748 (0.081)	-3.045 (1.385)	0.750 (0.081)
$\ln YS_i$	-0.2492 (0.583)	0.112 (0.238)	0.147 (0.037)	0.168 (0.038)
$(\ln HS_i)^2$	0.5509 (0.254)	****	0.622 (0.227)	****
$(\ln YS_i)^2$	0.0128 (0.027)	0.006 (0.026)	****	****
$(\ln HS_i)(\ln YS_i)$	0.0920 (0.143)	****	****	****
DC_i	0.0946 (0.043)	0.067 (0.043)	0.096 (0.043)	0.066 (0.0428)
$RSS =$	2.5918	2.8413	2.6067	2.8432
$TSS =$	8.0176	8.0176	8.0176	8.0176

*Nota: El símbolo "****" significa que la variable explicativa correspondiente se ha omitido.*

Todas los contrastes de hipótesis e intervalos de confianza están al nivel de significación del 5% y un nivel de confianza del 95%, respectivamente. Debe indicar claramente la hipótesis nula y alternativa, el estadístico de prueba, la región crítica (regla de rechazo) y la conclusión del contraste.

a. *Compare la bondad del ajuste de las cuatro ecuaciones penalizando por el número de variables explicativas incluidas. ¿Cuál de las cuatro ecuaciones de regresión estimadas proporciona el mejor ajuste? ¿Cuál proporciona el peor ajuste?*

Se debe emplear el coeficiente de determinación ajustado, R_a^2,

$$R_a^2 = 1 - \frac{SCR}{SCT}\frac{n-1}{n-k-1}$$

de tal forma que

$$\begin{aligned}
\text{Modelo (1)} &: \quad R_a^2 = 1 - \frac{2.5918}{8.0176}\frac{88-1}{88-6-1} = 0.652\,79 \\
\text{Modelo (2)} &: \quad R_a^2 = 1 - \frac{2.8413}{8.0176}\frac{88-1}{88-4-1} = 0.628\,54 \\
\text{Modelo (3)} &: \quad R_a^2 = 1 - \frac{2.6067}{8.0176}\frac{88-1}{88-4-1} = 0.659\,21 \\
\text{Modelo (4)} &: \quad R_a^2 = 1 - \frac{2.8432}{8.0176}\frac{88-1}{88-3-1} = 0.632\,72
\end{aligned}$$

El Modelo (3) proporciona el mejor ajuste y el (2) el peor.

b. *¿Cuál es el efecto estimado en el precio de la vivienda de un aumento del tamaño de la parcela de 500 pies cuadrados en casas de estilo colonial, con un tamaño de 500 pies cuadrados y una parcela de 1000 pies cuadrados? A continuación, compruebe si el tamaño de la parcela contribuye significativamente a ese efecto.*

Usando el modelo se obtiene

$$\Delta E\left[\ln P_i|\, \mathbf{X}\right] = \beta_2 \Delta \ln YS_i + \beta_4 \Delta \left(\ln YS_i\right)^2 + \beta_5 \left(\Delta \ln YS_i\right)\left(\ln HS_i\right)$$

Tenemos que $YS = 10$ pasa a tener un valor de $YS = 15$, por lo que

$$\begin{aligned}
\Delta \ln YS &= \ln 15 - \ln 10 = \ln 15/10 = 0.405 \\
\Delta \left(\ln YS_i\right)^2 &= \left(\ln 15\right)^2 - \left(\ln 10\right)^2 = 2.032
\end{aligned}$$

Las aproximaciones

$$\begin{aligned}
\Delta \ln YS &\approx \frac{\Delta YS}{YS} = \frac{5}{10} = 0.5 \\
\Delta \left(\ln YS_i\right)^2 &\approx 2\left(\ln YS\right)\Delta \ln YS \approx 2\left(\ln HS\right)\frac{\Delta YS}{YS} = 2\left(\ln 5\right)\frac{5}{10} = 1.609
\end{aligned}$$

son bastante imprecisas.

Por tanto

$$\begin{aligned}
\Delta \widehat{E\left[\ln P_i \right| \mathbf{X}]} &= -0.2492 * 0.405 + 0.0128 * 2.032 + 0.0920 * 2.032 * (\ln 5) \\
&= 0.226,
\end{aligned}$$

es decir que la casa valdrá un 22,6% más.

Para ver si el tamaño de la parcela es significativo en ese efecto, se realiza el contraste

$$\begin{aligned}
H_0 &: \beta_5 = 0 \\
H_1 &: \beta_5 \neq 0
\end{aligned}$$

Para ello se realiza un contraste de la t,

$$t_{\beta_5} = \frac{\hat{\beta}_5}{ee\left(\hat{\beta}_5\right)} = \frac{0.0920}{0.143} = 0.64336$$

que no es significativo comparado con el valor crítico al 5% de una normal estándar (1.96).

c. *Contraste que la elasticidad precio respecto a HS es constante.*

Usando la discusión en (b) se ha de contrastar

$$\begin{aligned}
H_0 &: \beta_3 = \beta_5 = 0 \\
H_1 &: \beta_3 \neq 0 \ \text{ y/o } \beta_5 \neq 0
\end{aligned}$$

ya que bajo H_0 la elasticidad es aproximadamente β_1 (constante). Para hacer el contraste se realiza un contraste de la F bajo homocedasticidad comparando el modelo (1) = no restringido y el modelo (2) = restringido

$$\begin{aligned}
F &= \frac{SCR_r - SCR_{nor}}{SCR_{nor}} \frac{n-k-1}{2} \\
&= \frac{2.8413 - 2.5918}{2.5918} \frac{88-6-1}{2} = 40.4
\end{aligned}$$

que comparado con el valor crítico al 5% de una $\chi^2_{2(0.05)}/2 \approx 3$, es muy significativo y se rechaza la hipótesis nula: la elasticidad respecto a HS no es constante (y depende del nivel de HS y/o YS).

d. *Contraste conjuntamente que ambas, la elasticidad precio respecto a HS y respecto a YS, son constantes.*

Se debe contrastar ahora

$$\begin{aligned}
H_0 &: \beta_3 = \beta_4 = \beta_5 = 0 \\
H_1 &: \quad H_0 \text{ es falsa}
\end{aligned}$$

ya que bajo H_0 la elasticidad de HS es aproximadamente β_1 (constante) y la de YS es β_2. Para hacer el contraste se realiza un contraste de la F bajo homocedasticidad comparando el modelo (1) = no restringido y el modelo (4) = restringido

$$\begin{aligned} F &= \frac{SCR_r - SCR_{nor}}{SCR_{nor}} \frac{n-k-1}{3} \\ &= \frac{2.8432 - 2.5918}{2.5918} \frac{88-6-1}{3} = 2.6190 \end{aligned}$$

que comparado con el valor crítico de una $\chi^2_{3(0.05)}/3 = 7.81/3 \approx 2.60$, es muy significativo y se rechaza la hipótesis nula: las elasticidades respecto a HS y YS no son constantes (alguna o las dos dependen del nivel de HS y/o YS).

2. *El gobierno de un país en desarrollo quiere implementar un programa donde las familias pobres reciben cupones que pueden ser utilizados para la compra de alimentos preenvasados con un alto valor nutricional. El gobierno decide establecer un experimento donde 500 familias (cada una con 1 niño) son asignados al azar a un grupo de tratamiento (elegible para cupones de alimentos, $T_i = 1$) y a un grupo de control (no elegible para recibir cupones de alimentos, $T_i = 0$). El gobierno ha contratado a una investigadora para estudiar el efecto de los cupones de alimentos sobre la probabilidad de que un niño tenga problemas de salud.*

Después del experimento, la investigadora estim un modelo por MCO donde la variable dependiente es H_i (una variable binaria que es igual a 1 si un niño tiene problemas de salud) y la variable explicativa es F_i (una variable binaria que vale 1 si una familia recibe cupones de alimentos). La siguiente tabla ofrece intervalos de confianza (I.C.) al 95% para los dos coeficientes (ambos utilizados errores estándar robustos), donde no se proporciona la estimación del coeficiente de la pendiente ni su error estándar.

I.C. del ajuste de H sobre F		
	I.C. al 95%	
Término constante	0.5788	0.7288
F	-0.3007	-0.1173

a. *Obtener la estimación del coeficiente de pendiente y su error estándar robusto a partir de la información disponible. A continuación, interpretar los coeficientes estimados.*

Usando que el IC al 95% es igual a

$$\hat{\beta}_F \pm 1.96 * ee\left(\hat{\beta}_F\right)$$

tenemos que

$$
\begin{aligned}
\hat{\beta}_F &= \frac{\left\{\hat{\beta}_F + 1.96 * ee\left(\hat{\beta}_F\right)\right\} + \left\{\hat{\beta}_F - 1.96 * ee\left(\hat{\beta}_F\right)\right\}}{2} \\
&= \frac{-0.1173 - 0.3007}{2} = -0.209 \\
ee\left(\hat{\beta}_F\right) &= \frac{\left\{\hat{\beta}_F + 1.96 * ee\left(\hat{\beta}_F\right)\right\} - \left\{\hat{\beta}_F - 1.96 * ee\left(\hat{\beta}_F\right)\right\}}{2 * 1.96} \\
&= \frac{-0.1173 + 0.3007}{2 * 1.96} = 0.047
\end{aligned}
$$

y se interpreta como una probabilidad 0.209 inferior de tener problemas de salud cuando se reciben los cupones de alimentos, todo lo demás igual.

b. *La investigadora encuentra que algunas de las familias del grupo control recibieron cupones de alimentos. Explique si podemos interpretar el coeficiente estimado MCO en F como el efecto marginal de cupones de alimentos en la salud infantil.*

En este caso no estamos seguros de que los otros factores contenidos en el error de un Modelo Lineal de Probabilidad permanezcan constantes entre las familias que recibieron los cupones y los que no (que sí era cierto respecto a la asignación aleatorio a los grupos de control y tratamiento), y por tanto puede haber endogeneidad y los estimadores MCO pueden estar sesgados.

Cuando la investigadora se entera de que algunas de las familias del grupo control recibieron cupones de alimentos, decide estimar el efecto de los cupones de alimentos utilizando un enfoque de variables instrumentales. Ella usa la asignación al grupo de tratamiento como instrumento para la recepción efectiva de los cupones de alimentos y obtiene los siguientes resultados de las estimaciones de la primera etapa.

Ajuste MCO de F sobre T		
	E.E. Robusto	Coeficiente
Término constante	0.307	0.376
T	0.307	0.624

c. *¿Cree que la condición de relevancia del instrumento se cumple? ¿Es T un instrumento débil? ¿Usted piensa que la condición de exogeneidad del instrumento se cumple?*

Relevancia: hay que contrastar si el coefiente de T es significativo:

$$
\begin{aligned}
H_0 &: \beta_T = 0 \\
H_1 &: \beta_T \neq 0
\end{aligned}
$$

mediante un contraste de la t,

$$
t = \frac{\hat{\beta}_T}{ee\left(\hat{\beta}_T\right)} = \frac{0.624}{0.307} = 2.0326
$$

por lo que es significativo al 5% y por tanto relevante. Para la debilidad computamos $F = t^2 = 2.0326^2 = 4.131\,5$, que es menor que 10, por lo que el instrumento es débil.

Exogeneidad: la asignación al grupo de control o tratamiento es aleatoria y por tanto independiente de cualquier otro factor que pueda afectar al estado de salud, y por tanto F es exógeno.

d. *La siguiente tabla muestra las medias de H y F para los asignados al grupo de tratamiento ($T = 1$) y para los asignados al grupo de control ($T = 0$). Utilizar los resultados de la tabla siguiente para obtener la estimación de variables instrumentales del efecto de los cupones de alimentos sobre la probabilidad de que un niño tenga problemas de salud.*

Estimadores de las medias condicionales		
	$T = 1$	$T = 0$
$\hat{E}(H\vert T)$	0.477	0.544
$\hat{E}(F\vert T)$	1	0.376

donde $\hat{E}(H|T = x)$ es la media muestral de los $H_i's$ que cumplen $T_i = x$, y $\hat{E}(F|T = x)$ es la media muestral de los $F_i's$ que cumplen $T_i = x$. Pista: puede usar el hecho de que

$$\widehat{Cov(H,T)} = \frac{1}{n}\sum_{i=1}^{n} H_i T_i - \bar{H}\bar{T} = \frac{n_1 n_0}{n}\left[\hat{E}(H|T=1) - \hat{E}(H|T=0)\right]$$

y

$$\widehat{Cov(F,T)} = \frac{1}{n}\sum_{i=1}^{n} F_i T_i - \bar{F}\bar{T} = \frac{n_1 n_0}{n}\left[\hat{E}(F|T=1) - \hat{E}(F|T=0)\right],$$

donde $n_1 = \sum_{i=1}^{500} T_i$ y $n_0 = 500 - n_1$.

Tenemos que

$$\hat{\beta}_F^{VI} = \frac{\widehat{Cov(H,T)}}{\widehat{Cov(F,T)}} = \frac{\frac{n_1 n_0}{n}\left[\hat{E}(H|T=1) - \hat{E}(H|T=0)\right]}{\frac{n_1 n_0}{n}\left[\hat{E}(F|T=1) - \hat{E}(F|T=0)\right]} = \frac{0.477 - 0.544}{1 - 0.376} = -0.1074$$

3. *Considere un modelo que relaciona el logaritmo de los salarios por hora en euros ($lnWage$), $Tenure$ (número de años de trabajo en la misma empresa), el género (usando la variable ficticia Man que toma el valor 1 si es hombre y 0 en caso contrario), y la educación,* *resumida por las siguientes variables explicativas,*

$$St0 = \begin{cases} 1 \text{ si solo Escuela Elemental} \\ 0 \text{ en otro caso.} \end{cases}$$
$$St1 = \begin{cases} 1 \text{ si acabada la Escuela Secundaria (High School)} \\ 0 \text{ en otro caso} \end{cases}$$
$$St2 = \begin{cases} 1 \text{ si Estudios Universitarios de Grado.} \\ 0 \text{ en otro caso} \end{cases}$$
$$St3 = \begin{cases} 1 \text{ si Estudios Universitarios de Postgrado.} \\ 0 \text{ en otro caso} \end{cases} .$$

El modelo

$$\ln(Wage) = \beta_0 + \beta_1 Man + \beta_2 Tenure + \beta_3 Tenure^2 + \beta_4 St1 + \beta_5 St2 + \beta_6 St3 + U$$

se estima por mínimos cuadrados ordinarios utilizando la Encuesta de Presupuestos Familiares española. La salida de un software econométrico estándar se proporciona a continuación.

Dependent Variable: $\ln(Wage)$				
Method: Least squares				
Included Observations: 27629				
	Coefficient	Std. Error	t-statistic	P-value
Man	0.280456	0.007082	39.59956	0.000
$Tenure$	0.087021	0.001114	78.13693	0.000
$Tenure^2$	-0.001697	$3.39E-05$	-50.07223	0.000
$St1$	0.086134	0.021001	4.101352	0.000
$St2$	0.217554	0.019959	10.90003	0.000
$St3$	0.752181	0.020249	37.14699	0.000
C	1.430037	0.019946	71.69381	0.000
R-squared	0.419248	Mean dependent var		2.387066
Adjusted R-squared	0.419122	S.D. dependent var		0.759744
S.E. of regression	0.579041	F-statistic		3323.418
Sum squared resid	9261.353	Prob(F-statistic)		0.0000

Se sabe que $\widehat{Cov}(\hat{\beta}_2, \hat{\beta}_3) = 0.000121$ *y* $\widehat{Cov}(\hat{\beta}_5, \hat{\beta}_6) = 0.000373$. *Se supone que todos los supuestos clásicos se cumplen. Todas los contrastes de hipótesis e intervalos de confianza están al nivel de significación del 5% y un nivel de confianza del 95%.*

a. *Proporcione las estimaciones MCO de los coeficientes* γ *en el modelo*

$$\ln(Wage) = \gamma_0 + \gamma_1 Man + \gamma_2 Tenure + \gamma_3 Tenure^2 + \gamma_4 St0 + \gamma_5 St1 + \gamma_6 St2 + U$$

Sustituyendo en

$$\ln(Wage) = \beta_0 + \beta_1 Man + \beta_2 Tenure + \beta_3 Tenure^2 + \beta_4 St1 + \beta_5 St2 + \beta_6 St3 + U$$

la relación

$$St3 = 1 - St0 - St1 - St2$$

se obtiene que

$$\begin{aligned}
& \ln(Wage) \\
= \; & \beta_0 + \beta_1 Man + \beta_2 Tenure + \beta_3 Tenure^2 + \beta_4 St1 + \beta_5 St2 + \beta_6 \{1 - St0 - St1 - St2\} + U \\
= \; & \{\beta_0 - \beta_6\} + \beta_1 Man + \beta_2 Tenure + \beta_3 Tenure^2 - \beta_6 St0 + \{\beta_4 - \beta_6\} St1 + \{\beta_5 - \beta_6\} St2
\end{aligned}$$

y por tanto

$$\begin{aligned}
\hat{\gamma}_0 &= \hat{\beta}_0 - \hat{\beta}_6 = 1.430037 - 0.752181 = 0.677\,86 \\
\hat{\gamma}_1 &= \hat{\beta}_1 = 0.280456 \\
\hat{\gamma}_2 &= \hat{\beta}_2 = 0.087021 \\
\hat{\gamma}_3 &= \hat{\beta}_3 = 0.001697 \\
\hat{\gamma}_4 &= -\hat{\beta}_6 = -0.752181 \\
\hat{\gamma}_5 &= \hat{\beta}_4 - \hat{\beta}_6 = 0.086134 - 0.752181 = -0.666\,05 \\
\hat{\gamma}_6 &= \hat{\beta}_5 - \hat{\beta}_6 = 0.217554 - 0.752181 = -0.534\,63
\end{aligned}$$

b. *Contraste que los salarios de los trabajadores con estudios universitarios de postgrado son más altos que los salarios que los de los trabajadores con estudios universitarios de grado.*

Contraste de

$$\begin{aligned}
H_0 &: \beta_6 - \beta_5 = 0 \\
H_0 &: \beta_6 - \beta_5 > 0
\end{aligned}$$

que equivale a

$$\begin{aligned}
H_0 &: \gamma_6 = 0 \\
H_0 &: \gamma_6 < 0
\end{aligned}$$

mediante el contraste de la t

$$\begin{aligned}
t &= \frac{\hat{\beta}_6 - \hat{\beta}_5}{ee\left(\hat{\beta}_6 - \hat{\beta}_5\right)} = \frac{\hat{\beta}_6 - \hat{\beta}_5}{\left(\widehat{Var}\left(\hat{\beta}_6\right) + \widehat{Var}\left(\hat{\beta}_5\right)^2 - 2\widehat{Cov}(\hat{\beta}_5, \hat{\beta}_6)\right)^{1/2}} \\
&= \frac{\hat{\beta}_6 - \hat{\beta}_5}{\left(ee\left(\hat{\beta}_6\right)^2 + ee\left(\hat{\beta}_5\right)^2 - 2\widehat{Cov}(\hat{\beta}_5, \hat{\beta}_6)\right)^{1/2}} \\
&= \frac{0.752181 - 0.217554}{(0.020249^2 + 0.019959^2 - 2 * 0.000373)^{1/2}} = 67.\,689,
\end{aligned}$$

que es significativo y por tanto se rechaza la nula en favor de la alternativa de diferencia positiva.

c. *Contraste que tenure no tiene efecto sobre los salarios.*

Se ha de contrastar

$$\begin{aligned} H_0 &: \beta_2 = \beta_3 = 0 \\ H_1 &: \beta_2 \neq 0 \text{ y/o } \beta_3 \neq 0 \end{aligned}$$

Se utiliza el estadístico

$$\begin{aligned} F &= \left(\hat{\beta}_2, \hat{\beta}_3\right)' \begin{bmatrix} \widehat{Var}\left(\hat{\beta}_2\right) & \widehat{Cov}(\hat{\beta}_2, \hat{\beta}_3) \\ \widehat{Cov}(\hat{\beta}_2, \hat{\beta}_3) & \widehat{Var}\left(\hat{\beta}_3\right) \end{bmatrix}^{-1} \begin{pmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} \Big/ 2 \\ &= (0.087021, 0.001697)' \begin{bmatrix} 0.001114^2 & 0.000121 \\ 0.000121 & (3.39E-05)^2 \end{bmatrix}^{-1} \begin{pmatrix} 0.087021 \\ 0.001697 \end{pmatrix} \Big/ 2 \\ &= 2.5501 < \chi^2_{2,(0.05)} \big/ 2 = 5.99/2 \approx 3 \end{aligned}$$

Por tanto, no rechazamos H_0.

d. *Si sospecha que hay variables no observables, como "habilidad", correlacionados con los salarios y con la educación, explicar cómo se estimaría el modelo por mínimos cuadrados en dos etapas. Explique cuántos instrumentos son necesarios y proporcione ejemplos de posibles instrumentos.*

MC2E: 1^a etapa: regresar $St1$, $St2$, $St3$ sobre las variables exógenas ($cons\tan te$, Man, $Tenure$, $Tenure^2$, y al menos tres instrumentos adicionales para $St1$, $St2$, $St3$).

2^a etapa, estimar la ecuación original sustituyendo $St1$, $St2$, $St3$ por la predicciones MCO de la primera etapa.

Se necesitan al menos 3 instrumentos, por ejemplo: educación de la madre, del padre, ...

4. *Los siguientes modelos han sido estimados por MCO para explicar el consumo de alcohol (A), medido como número de bebidas por semana, utilizando una muestra aleatoria de 500 personas mayores de 22 años (errores estándar robustos a la heterocedasticidad entre paréntesis).*

$$\hat{A}_i = \underset{(1.1)}{2.4} - \underset{(0.4)}{0.8} Educ_i - \underset{(0.1)}{0.3} IQ_i \tag{1}$$

donde $Educ_i$ y IQ_i son años de educación y el cociente de inteligencia del individuo i-ésimo, respectivamente, que se asumen exógenos a menos que se indique lo contrario. La covarianza muestral entre Educ y IQ es $\widehat{Cov}(Educ, IQ) = 2.7$.

a. *Suponga que ha ignorado IQ y ha estimado el modelo*

$$A_i = \beta_0 + \beta_1 Educ_i + U_i. \tag{2}$$

Proporcione una expresión para el sesgo del estimador MCO del efecto marginal de Educ en este modelo, y explique cual es el signo esperado del sesgo.

La expresión para el sesgo del estimador MCO en el modelo con variable omitida $\tilde{\beta}_1$ es

$$E\left[\tilde{\beta}_1\right] - \beta_1 = \beta_2 \frac{Cov\,(Educ, IQ)}{Var\,(Educ)}$$

y pensando que $Cov\,(Educ, IQ) \approx \widehat{Cov}(Educ, IQ) = 2.7 > 0$, $Var\,(Educ) > 0$ y que $\beta_2 < 0$, entonces el sesgo será negativo: el valor de $\tilde{\beta}_1$ será sistemáticamente menor que el valor verdadero de β_1.

b. *Suponga que la estimación del modelo con variable omitida es*

$$\hat{A}_i = \underset{(1.1)}{2.4} - \underset{(0.4)}{1.2} Educ_i$$

Proporcione una expresión para la varianza muestral de Educ.

La relación entre la estimación del modelo con variable omitida $\tilde{\beta}_1$ y la del modelo completo, $\hat{\beta}_1$ es

$$\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \frac{\widehat{Cov}\,(Educ, IQ)}{\widehat{Var}\,(Educ)}$$

y, por tanto, sustituyendo,

$$-1.2 = -0.8 - 0.3 \frac{2.7}{\widehat{Var}\,(Educ)}$$

y finalmente

$$\widehat{Var}\,(Educ) = \frac{-0.3 * 2.7}{-1.2 + 0.8} = 2.025.$$

c. *Se sabe que "madurez intelectual" (IM) puede explicar A_i mejor que IQ_i, pero no se observa. A continuación, se propone estimar el modelo (2) por mínimos cuadrados en dos etapas, utilizando como variables instrumentales "educación del padre" $Fath - educ$ y "educación de la madre" $Moth - educ$. Explique como contrastaría que estos dos instrumentos son conjuntamente exógenos.*

Se llevaría a cabo un contraste de restricciones de sobreidentificación (Sargan-Hansen), regresando los residuos del modelo ajustado por MC2E sobre todas las variables exógenas y se computaría el estadístico $J = mF$, donde $m = 2$ es el número de instrumentos y F es el valor del estadístico F del contraste conjunto de significación de los m instrumentos. El estadístico J se compararía con el valor crítico de una $\chi^2_{m-k} = \chi^2_1$ donde $k = 1$ es el número de regresores endógenos, $Educ$ en este caso. Si es significativo se rechaza la nula de exogeneidad conjunta de los dos instrumentos.

Valores Críticos:
$\chi^2_{1,(0.05)} = 3.84$, $\chi^2_{2,(0.05)} = 5.99$, $\chi^2_{3,(0.05)} = 7.81$, $\quad Z_{(0.025)} = 1.96$, $Z_{(0.05)} = 1.65$.