EXAMEN FINAL DE ECONOMETRÍA

SOLUCIONES

Conteste cada pregunta en un cuadernillo diferente en dos horas y media

1. *Sean (Y, X, W) tres variables aleatorias relacionadas por el siguiente modelo de regresión lineal*

$$Y = \beta_0 + \beta_1 X + \beta_2 W + u, \tag{1}$$

donde se sabe que

$$\mathbb{E}(u|X, W) = \mathbb{E}(u|W) = \gamma_0 + \gamma_2 W. \tag{2}$$

(a) *Escriba la ecuación (2) como un modelo de regresión lineal para explicar la variable u en función de W donde deberá demostrar que el nuevo término de error v satisface $E(v|X, W) = 0$. Sustituya ese modelo para u en (1) e interprete los coeficientes del modelo obtenido que relaciona Y con X, W y v.*

Se puede escribir

$$u = \gamma_0 + \gamma_2 W + v$$

donde el término de error $v := u - (\gamma_0 + \gamma_2 W) = u - \mathbb{E}(u|W) = u - \mathbb{E}(u|X, W)$, usando *(2)* para las dos igualdades, satisface

$$\begin{aligned}
\mathbb{E}(v|X, W) &= \mathbb{E}(u - \mathbb{E}(u|X, W)|X, W) \quad \text{definición de } u \\
&= \mathbb{E}(u|X, W) - \mathbb{E}(\mathbb{E}(u|X, W)|X, W) \quad \text{linealidad de } \mathbb{E}(\cdot|X, W) \\
&= \mathbb{E}(u|X, W) - \mathbb{E}(u|X, W) \quad \text{ley de esperanzas iteradas} \\
&= 0.
\end{aligned}$$

[5 puntos]

Sustituyendo en (1) se obtiene

$$\begin{aligned}
Y &= \beta_0 + \beta_1 X + \beta_2 W + u \\
&= \beta_0 + \beta_1 X + \beta_2 W + (\gamma_0 + \gamma_2 W + v) \\
&= (\beta_0 + \gamma_0) + \beta_1 X + (\beta_2 + \gamma_2) W + v, \qquad (3)
\end{aligned}$$

donde error v satisface el supuesto de regresión múltiple $\mathbb{E}(v|X, W) = 0$, y por tanto el coeficiente de X, β_1 mide el efecto causal de X (manteniendo constante W). El coeficiente de W, $\beta_2 + \gamma_2$, también mide el efecto causal de W (manteniendo constante X) pero en general es diferente de β_2, que por tanto no mide el efecto causal de W. El término constante también es diferente respecto al modelo (1). [5 puntos]

(b) *Se observa una muestra aleatoria simple $\{Y_i, X_i, W_i\}_{i=1}^n$ de (Y, X, W) y se considera el estimador de mínimos cuadrados ordinarios (MCO) de los coeficientes en el modelo (1) donde Y es la variable explicada y (X, W) las variables explicativas. Justifique si este estimador es sesgado o insesgado para los parámetros $(\beta_0, \beta_1, \beta_2)$ cuando se cumple (2). ¿Qué ocurriría si en la regresión no se incluye el regresor W?*

El estimador de MCO estimará consistentemente los parámetros del modelo cuyo término de error tenga las propiedades habituales, es decir el modelo *(3)*, y por tanto estimará consistentemente el coeficiente de X, β_1, pero estimará inconsistentemente los parámetros β_0 y β_1, ya que los coeficientes de la constante y de W son $\beta_0 + \gamma_0 \neq \beta_0$ y $\beta_2 + \gamma_2 \neq \beta_2$, respectivamente. [5 puntos]

Si no se introduce W, la regresión tendrá un sesgo por variables omitidas si existe correlación entre X y W, y el estimador MCO de β_1 también estará sesgado y será inconsistente. [5 puntos]

(c) *Un estudio posterior confirma que la condición (2) no es sostenible, ya que en realidad se cumple que*

$$\mathbb{E}(u|X,W) = \alpha_0 + \alpha_1 X + \alpha_2 W, \quad \alpha_1 \neq 0,$$

pero se dispone de datos de otra variable Z que satisface

$$\mathbb{E}(u|Z,W) = \mathbb{E}(u|W) = \delta_0 + \delta_2 W, \quad \delta_2 \neq 0. \tag{4}$$

*¿Cuál sería ahora su estrategia de estimación del modelo (1) con las restricciones (**??**)-(4) ? ¿Qué coeficientes de ese modelo se podrían estimar consistentemente? Justifique su respuesta.*

Dada la nueva condición sobre $\mathbb{E}(u|X,W)$, reproduciendo el análisis anterior en (b) se puede obtener que

$$Y = (\beta_0 + \alpha_0) + (\beta_1 + \alpha_1)X + (\beta_2 + \alpha_2)W + v$$

por lo que los tres coeficientes $(\beta_0, \beta_1, \beta_2)$ se estimarán inconsistentemente por MCO. Sin embargo, usando la condición (*4*) se puede emplear Z como variable instrumental para estimar la ecuación por MC2E. [5 puntos]

Como en (b), no se podrían estimar consistentemente los parámetros β_1 y β_2, sólo el coeficiente de X, β_1, como se puede comprobar escribiendo el modelo con un nuevo término de error

$$Y = (\beta_0 + \delta_0) + \beta_1 X + (\beta_2 + \delta_2)W + m$$

para el que se cumple $\mathbb{E}(m|Z,W) = 0$. [5 puntos]

2. *Utilizando datos de sección cruzada en Estados Unidos para todos los estados, se quiere estudiar la relación entre el consumo de cannabis y los ingresos de los individuos. Con este fin se considera el sistema de ecuaciones*

$$\log(ingresos) = \beta_0 + \beta_1 cannabis + u_1 \tag{5}$$

$$cannabis = \gamma_0 + \gamma_1 \log(ingresos) + \gamma_2 multa + \gamma_3 prisión + u_2, \tag{6}$$

donde cannabis es el consumo mensual de cannabis, multa es la multa estándar impuesta por posesión de cannabis en el estado de residencia y prisión es una variable binaria que toma el valor 1 si en el estado de residencia se puede condenar a un individuo a prisión por posesión de cannabis para consumo personal. Se supone que tanto multa como prisión varían entre estados y $\gamma_1 \neq 0$.

(a) *Si las variables multa y prisión son exógenas (no están correlacionadas con ningún término de error), encuentre la forma reducida de la variable cannabis y demuestre que cannabis es endógena en la ecuación (5).*

Reemplanzando la ecuación (5) en (6) y agrupando términos se obtiene

$$\begin{aligned} cannabis &= \gamma_0 + \gamma_1 \log(ingresos) + \gamma_2 multa + \gamma_3 prisión + u_2 \\ &= \gamma_0 + \gamma_1 (\beta_0 + \beta_1 cannabis + u_1) + \gamma_2 multa + \gamma_3 prisión + u_2 \\ &= (\gamma_0 + \gamma_1 \beta_0) + \gamma_1 \beta_1 cannabis + \gamma_2 multa + \gamma_3 prisión + u_2 + \gamma_1 u_1 \\ &= \frac{\gamma_0 + \gamma_1 \beta_0}{1 - \gamma_1 \beta_1} + \frac{\gamma_2}{1 - \gamma_1 \beta_1} multa + \frac{\gamma_3}{1 - \gamma_1 \beta_1} prisión + \frac{u_2 + \gamma_1 u_1}{1 - \gamma_1 \beta_1} \\ &= \pi_0 + \pi_1 multa + \pi_2 prisión + v_2 \end{aligned} \tag{7}$$

que relaciona *cannabis* con las dos variables exógenas del modelo y un término de error que es una combinación lineal de los errores de las dos ecuaciones de la forma estructural. [5 puntos]

Por tanto, usando la exogeneidad de *multa* y *prisión*, se puede calcular

$$\begin{aligned} Cov\,(cannabis, u_1) &= Cov\,(\pi_0 + \pi_1 multa + \pi_2 prisi\acute{o}n + v_2, u_1) \\ &= Cov\,(v_2, u_1) = Cov\left(\frac{u_2 + \gamma_1 u_1}{1 - \gamma_1 \beta_1}, u_1\right) \\ &= \frac{Cov\,(u_2, u_1) + \gamma_1 Var\,(u_1)}{1 - \gamma_1 \beta_1}, \end{aligned}$$

que es diferente de cero porque $\gamma_1 \neq 0$ y $Var\,(u_1) > 0$ (aunque $Cov\,(u_2, u_1) = 0$), lo que demuestra que *cannnabis* es endógena porque está correlada con el término de error u_1 de la ecuación (5). [5 puntos]

(b) *¿Qué restricciones deben cumplir los coeficientes $\gamma's$ en la ecuación (6) para poder estimar sin sesgo los coeficientes $\beta's$ en la ecuación (5)? Explique si sería posible contrastar esas restricciones, directa o indirectamente.*

Para poder estimar consistentemente la ecuación (5) debemos disponer de instrumentos válidos para *cannabis* para usar MC2E: *multa* y *prisión* son exógenos, pero hay que comprobar si son relevantes. Para ello es necesario que en la forma reducida (7) de *cannabis* al menos uno de los dos coeficientes, π_1 ó π_2, de *multa* y *prisión* respectivamente, sean diferentes de cero. Como

$$\pi_1 = \frac{\gamma_2}{1 - \gamma_1 \beta_1} \quad \text{y} \quad \pi_2 = \frac{\gamma_3}{1 - \gamma_1 \beta_1}$$

eso es verdad si y sólo si $\gamma_2 \neq 0$ y/o $\gamma_3 \neq 0$. [5 puntos]

En principio los estimadores MCO de la ecuación (6) y los parámetros γ no son consistentes, pero sí se puede estimar consistentemente la forma reducida (7) de *cannabis* y los parámetros $\pi's$ por MCO, y con esas estimaciones contrastar la hipótesis nula

$$H_0 : \pi_1 = \pi_2 = 0$$

en contra de la alternativa $H_1 : \pi_1 \neq 0$ y/o $\pi_2 \neq 0$ con un contraste habitual de la F. [5 puntos]

(c) *Explique cómo estimaría los coeficientes $\beta's$ de la ecuación (5) y cómo obtendría errores estándard válidos bajo heterocedasticidad.*

Se usaría MC2E:

1^a etapa: estimación de la forma reducida (7) de *cannabis* por MCO, y obtención de las predicciones

$$\widehat{cannabis} = \hat{\pi}_0 + \hat{\pi}_1 multa + \hat{\pi}_2 prisi\acute{o}n.$$

2^a etapa: estimación de la ecuación (5) por MCO reemplazando *cannabis* por $\widehat{cannabis}$,

$$\log(ingresos) = \beta_0 + \beta_1 \widehat{cannabis} + error.$$

[5 puntos]

Los errores estándard de esta regresión no son los correctos porque no tienen en cuenta que $\widehat{cannabis}$ es una variable estimada y deberían usar una fórmula que tenga en cuenta la heterocedasticidad conditional de u_1 (respecto a los instrumentos usados en el ajuste MC2E). [5 puntos]

(d) *Explique cómo contrastaría que los instrumentos para estimar los parámetros $\beta's$ en la ecuación (5) son realmente exógenos; esto es, que ninguno está correlacionado con el término de error. En particular, explique como obtendría el estadístico del contraste y sus valores críticos o p-valores, para implementarlo.*

Para comprobar la exogeneidad de los instrumentos *multa* y *prisión*, hay que contrastar la hipótesis nula de que $Cov\,(multa, u_1) = Cov\,(prisi\acute{o}n, u_1) = 0$, que es este caso es posible ya que la ecuación está sobreidentificada (por que se dispone de dos instrumentos, *multa* y *prisión*, para un sólo regresor endógeno, *cannabis*). [3 puntos]

Para el contraste se usaría un contraste J de sobre-identificación que sigue este procedimiento:

1. Estimar la ecuación estructural (5) por MC2E y obtener las predicciones $\log(\widehat{ingresos})$,

$$\log(\widehat{ingresos}) = \hat{\beta}_0 + \hat{\beta}_1 cannabis$$

y los residuos $\hat{u}_1 = \log(ingresos) - \log(\widehat{ingresos})$.

2. Estimar por MCO la regresión auxiliar

$$\hat{u}_1 = \delta_0 + \delta_1 multa + \delta_2 prisi\acute{o}n + error$$

y calcular el estadístico F que se usaría para contrastar las hipótesis

$$\begin{aligned} H_0 &: \delta_1 = \delta_2 = 0 \\ H_1 &: H_0 \text{ es falsa.} \end{aligned}$$

3. Calcular el estadístico J

$$J = mF = 2F$$

donde $m = 2$ es el número de instrumentos. [4 puntos]

El estadístico J se compararía con el valor crítico de la distribución $\chi^2_{m-k} = \chi^2_1$ donde $m - k = 2 - 1 = 1$ son las condiciones de sobre-identificación y k el número de variables endógenas en la ecuación estructural original, rechazando la hipótesis de exogeneidad si el estadístico J es mayor que el correspondiente valor crítico, con p-valor igual a

$$p - valor = \Pr\left(\chi^2_1 > J\right).$$

[3 puntos]

3. *Se quiere investigar la función de demanda de subscripciones de revistas científicas por parte de las bibliotecas universitarias. Para ello se dispone de datos para 180 revistas de economía referentes al año 2000. Se considera el modelo de regresión lineal*

$$\log(subs) = \beta_0 + \beta_1 \log(ppcite) + \beta_2 (\log(ppcite))^2 + \beta_3 \log(age) + \beta_4 \log(ppcite) * \log(age) + u, \quad (8)$$

donde subs es el número de bibliotecas que están subscritas a una determinada revista, ppcite es el precio de la subscripción divido por el número total de citas de los artículos publicados en esa revista y age es la antigüedad de la revista en años. $\log(x)$ *es el logaritmo neperiano de* x. *El término de error u satisface los supuestos habituales del modelo de regresión múltiple clásico que garantizan que las estimaciones MCO son insesgadas..*

(a) *Proporcione una expresión en términos de los coeficientes del modelo (8) que aproxime la variación esperada en el número de subscripciones cuando el precio por cita aumenta un 1% para una revista de 20 años con* $ppcite = 10$. *Explique como contrastaría que dicho efecto es igual o mayor que* -1 *y bajo qué condiciones es válido el contraste propuesto.*

Se necesita calcular la elasticidad de demanda con la aproximación $(\Delta subs)/subs \approx \Delta \log(subs)$,

$$\begin{aligned} \frac{\partial \mathbb{E}\left(\log(subs) \mid ppcite = 10, age = 20\right)}{\partial \log(ppcite)} &= \beta_1 + 2\beta_2 \log(ppcite = 10) + \beta_4 \log(age = 20) \\ &= \beta_1 + 2\beta_2 \log(10) + \beta_4 \log(20), \end{aligned}$$

que proporciona el porcentaje de incremento en el número de subscripciones. [3 puntos]

Para hacer el contraste de

$$\begin{aligned} H_0 &: \beta_1 + 2\beta_2 \log(10) + \beta_4 \log(20) \geq -1 \\ H_1 &: \beta_1 + 2\beta_2 \log(10) + \beta_4 \log(20) < -1 \end{aligned}$$

se puede calcular el estadístico t robusto a la heterocedasticidad

$$t = \frac{\hat{\beta}_1 + 2\hat{\beta}_2 \log(10) + \hat{\beta}_4 \log(20) + 1}{ee\left(\hat{\beta}_1 + 2\hat{\beta}_2 \log(10) + \hat{\beta}_4 \log(20)\right)}$$

y compararlo con una normal estándard. Se rechazaría H_0 al 5% si $t < 1.645$. [4 puntos]
El error estándard $ee\left(\hat{\beta}_1 + 2\hat{\beta}_2 \log(10) + \hat{\beta}_4 \log(20)\right)$ es la raíz cuadrada de

$$\begin{aligned} & \widehat{Var}\left(\hat{\beta}_1 + 2\hat{\beta}_2 \log(10) + \hat{\beta}_4 \log(20)\right) \\ = & \widehat{Var}\left(\hat{\beta}_1\right) + 4\log(10)^2\, \widehat{Var}\left(\hat{\beta}_2\right) + \log(20)\, \widehat{Var}\left(\hat{\beta}_4\right) \\ & +4\log(10)\, \widehat{Cov}\left(\hat{\beta}_1, \hat{\beta}_2\right) + 2\log(20)\, \widehat{Cov}\left(\hat{\beta}_1, \hat{\beta}_4\right) + 4\log(20)\log(10)\, \widehat{Cov}\left(\hat{\beta}_2, \hat{\beta}_4\right), \end{aligned}$$

donde $\widehat{Var}$ y $\widehat{Cov}$ usan las fórmulas generales robustas a la heterocedasticidad. [3 puntos]
Alternativamente se puede construir un estadístico F no robusto a la heterocedasticidad usando el modelo restringido bajo H_0 que se obtiene sustituyendo la restricción en el modelo original,

$$\begin{aligned} \log(subs) &= \beta_0 + (-2\beta_2 \log(10) - \beta_4 \log(20) - 1)\log(ppcite) + \beta_2 (\log(ppcite))^2 \\ & +\beta_3 \log(age) + \beta_4 \log(ppcite) * \log(age) + u \end{aligned}$$

y despejando,

$$Y^* = \beta_0 + \beta_2\left((\log(ppcite))^2 - 2\log(10)\right) + \beta_3 \log(age) + \beta_4 \log(ppcite/20) * \log(age) + u,$$

donde $Y^* = \log(subs) + \log(ppcite)$ y

$$F = \frac{SCR_{restringido} - SCR_{no\ restringido}}{SCR_{no\ restringido}} \frac{n-k-1}{q}$$

no se puede calcular usando los R^2 porque cambia la variable dependiente del modelo restringido, $Y^* \neq \log(subs)$, y a continuación calcular $t = signo * \sqrt{F}$ donde el *signo* es el signo de $\hat{\beta}_1 + 2\hat{\beta}_2 \log(10) + \hat{\beta}_4 \log(20) + 1$ para hacer un contraste unilateral sólo válido bajo homocedasticidad [5 puntos, sólo si se comenta la validez usando SCR, no R^2, y sólo bajo homocedasticidad y se construye el contraste unilateral con el signo.]

(b) *Contraste si la elasticidad de la demanda de subscripciones a los precios depende de la antigüedad de la revista en el modelo (8) usando las estimaciones de la Tabla 1. ¿Bajo qué condiciones es este contraste válido?*

La elasticidad de la demanda de subscripciones es el parámetro calculado en (a), por lo que se require contrastar las hipótesis

$$\begin{aligned} H_0 &: \beta_4 = 0 \\ H_1 &: \beta_4 \neq 0, \end{aligned}$$

para lo que se usará el contraste de la t para $\beta_4 = \beta_{\log(ppcite)*\log(age)}$ [3 puntos]
Usando las estimaciones del Modelo (3) en la Tabla 1

$$t_4 = \frac{\hat{\beta}_4}{ee\left(\hat{\beta}_4\right)} = \frac{0.257}{0.106} = 2.4245$$

y comparado con el valor crítico de una normal estándard al 5%, 1.96, es significativo,

$$|t_4| > 1.96,$$

por lo que se rechaza la hipótesis nula, y se confirma que la elasticidad de demanda depende la antigüedad de la revista. [4 puntos]

Este contraste sólo es válido asintóticamente si los errores estándard son robustos a la heterocedasticidad, o si se asumen condiciones de homocedasticidad condicional. [3 puntos]

(c) *Alternativamente se decide estimar el siguiente modelo usando la variable binaria* $age40$ *que es igual a 1 si la antigüedad de la revista es igual o inferior a 40 años, y 0 en caso contrario,*

$$\begin{aligned}\log(subs) &= \beta_0 + \beta_1 \log(ppcite) + \beta_2 (\log(ppcite))^2 + \beta_3 age40 \\ &\quad + \beta_4 \log(ppcite) * age40 + \beta_5 (\log(ppcite))^2 * age40 + u.\end{aligned} \tag{9}$$

Contraste si la elasticidad de la demanda de subscripciones a los precios es constante para todos los niveles de precios en el modelo (9) usando las estimaciones de la Tabla 1. ¿Bajo qué condiciones es este contraste válido?

Ahora se puede comprobar que la elasticidad es

$$\frac{\partial \mathbb{E}\left(\log(subs)|\, ppcite, age\right)}{\partial \log(ppcite)} = \beta_1 + 2\beta_2 \log(ppcite) + \beta_4 age40 + 2\beta_5 \log(ppcite) * age40.$$

La elasticidad de la demanda de subscripciones depende del nivel de precios a través de los parámetros β_2 y β_5 por lo que se require contrastar las hipótesis

$$\begin{aligned} H_0 &: \beta_2 = \beta_5 = 0 \\ H_1 &: \beta_2 \neq 0 \text{ y/o } \beta_5 \neq 0, \end{aligned}$$

para lo que se usará un contraste de la F para $\beta_2 = \beta_{\log(ppcite)^2} = 0$ y $\beta_5 = \beta_{\log(ppcite)^2 * age40} = 0$ usando las estimaciones del Modelo (6) en la Tabla 1, comparadas con el Modelo (4), que es la regresión restringida bajo H_0. [4 puntos]

Con los datos disponibles, sólo se puede calcular el estadístico F no robusto a heterocedasticidad, que sólo será válido bajo homocedasticidad condicional [3 puntos, incluyendo 2 por la homocedasticidad condicional]

$$\begin{aligned} F &= \frac{R^2_{no\ restringido} - R^2_{restringido}}{1 - R^2_{no\ restringido}} \frac{n-k-1}{q} \\ &= \frac{0.633 - 0.632}{1 - 0.633} \frac{180-5-1}{2} = 0.237\,06 \end{aligned}$$

que comparado con el valor crítico al 5% de una $\chi^2_2/2 = 5.99/2 \approx 3$, no es significativo

$$F \leq 3,$$

y no se rechaza la hipótesis nula, por lo que no se puede rechazar que la elasticidad de demanda sea constante para todos los niveles de precios. [3 puntos]

ALGUNOS VALORES CRÍTICOS: $Z_{0.90} = 1.282$, $Z_{0.95} = 1.645$, $Z_{0.975} = 1.96$, $\chi^2_{2,95} = 5.99$, $\chi^2_{2,975} = 7.378$, $\chi^2_{3,95} = 7.81$, $\chi^2_{3,975} = 9.3484$, $\chi^2_{4,95} = 9.49$, $\chi^2_{4,975} = 11.1433$, $\chi^2_{4,95} = 9.49$, $\chi^2_{4,975} = 11.1433$, $\chi^2_{5,95} = 11.07$, $\chi^2_{5,975} = 12.8325$, donde $\mathbb{P}(Z \leq Z_\alpha) = \alpha$ y $\mathbb{P}\left(\chi^2_m \leq \chi^2_{m,\alpha}\right) = \alpha$, Z está distribuida como una normal com media cero y varianza uno, y χ^2_m como una chi-cuadrado con m grados de libertad.

Table 1: Tabla de Regresiones

	(1)	(2)	(3)	(4)	(5)	(6)
Variable Dependiente:	log(*subs*)	log(*subs*)	log(*subs*)	log(*subs*)	log(*subs*)	log(*subs*)
log(*ppcite*)	-0.561	-0.460	-1.468	-0.382	-0.455	-0.351
	(0.0449)	(0.0446)	(0.385)	(0.0369)	(0.0474)	(0.0597)
$\log(ppcite)^2$	-0.0291		0.00721		-0.0142	0.0164
	(0.0177)		(0.0219)		(0.0170)	(0.0197)
log(*age*)		0.531	0.556			
		(0.222)	(0.247)			
$\log(ppcite) * \log(age)$			0.257			
			(0.106)			
*age*40				-0.534	-0.624	-0.491
				(0.117)	(0.126)	(0.146)
$\log(ppcite) * age40$				-0.205		-0.228
				(0.0805)		(0.0854)
$\log(ppcite)^2 * age40$						-0.0272
						(0.0415)
Constant	4.839	2.724	2.699	5.054	5.050	5.028
	(0.0680)	(0.841)	(0.934)	(0.0766)	(0.0827)	(0.0822)
Observaciones	180	180	180	180	180	180
R^2	0.564	0.584	0.607	0.632	0.620	0.633

Errores Estándard en paréntesis