EXAMEN FINAL DE ECONOMETRÍA

Conteste cada pregunta en un cuadernillo diferente en dos horas

1. Sean (Y, X, Z, W) cuatro variables aleatorias tales que,

$$\mathbb{E}(Y|X, Z, W) = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 W. \quad (1)$$

Solamente estamos interesados en estimar β_1 , los parámetros β_0 , β_2 y β_3 carecen de interés. Resulta que observamos una muestra aleatoria simple $\{Y_i, X_i, Z_i\}_{i=1}^n$ de (Y, X, Z) , pero no observamos W . Sin embargo, sabemos que,

$$\mathbb{E}(W|X, Z) = \gamma_0 + \gamma_1 Z. \quad (2)$$

- a. Expresé (1) utilizando un término de error U con media condicional, dados X, Z y W , igual a cero. Haga lo mismo con (2) en términos de un término de error V con media condicional, dados X y Z , igual a cero. Combine estas dos expresiones para expresar Y en términos de X, Z y un término de error $\epsilon = (U + \beta_3 V)$, donde los coeficientes dependen de $\beta_0, \beta_1, \beta_2, \beta_3, \gamma_0$ y γ_1 . ¿En cuanto se espera que variará Y cuando X aumenta en una unidad y solamente Z permanece constante? Justifique su respuesta.
- b. Considere el estimador de mínimos cuadrados ordinarios (MCO) de los coeficientes en el modelo donde Y es la variable explicada y (X, Z) las variables explicativas. En vista de su respuesta al apartado a., explique de forma razonada por qué el coeficiente MCO de X es un estimador sesgado o insesgado de β_1 . Haga lo mismo para el coeficiente MCO de Z con respecto a β_2 .
- c. ¿Cuál sería su respuesta al apartado b. si se sabe que W es independiente de (X, Z) ?
2. Considere la estimación de una función de producción que relaciona la cantidad producida (output), Y , con la cantidad utilizada de trabajo, L , y la cantidad utilizada de capital (instalaciones y maquinaria), K . Un modelo econométrico flexible es el basado en la función translog,

$$\log Y = \beta_0 + \beta_1 \log L + \beta_2 \log K + \beta_3 \left(\frac{1}{2} \log^2 L \right) + \beta_4 \left(\frac{1}{2} \log^2 K \right) + \beta_5 (\log K \cdot \log L) + U. \quad (3)$$

Utilizando observaciones en 27 estados de Estados Unidos en el periodo 1957-1958 sobre la industria de metal primario (SC33), los cuales se han utilizado en un buen número de estudios en economía industrial, se ha estimado esta función por mínimos cuadrados ordinarios. Suponemos, en lo que sigue, que se cumplen todos los supuestos clásicos.

Una función de producción más restrictiva, que impone una elasticidad de sustitución constante, es la Cobb-Douglas,

$$\log Y = \beta_0 + \beta_1 \log L + \beta_2 \log K + U. \quad (4)$$

Las siguientes salidas de GRET (Modelos 1-5) estiman modelo (3), (4) y algunas variaciones de los mismos. Los errores estándar son robustos frente a heterocedasticidad.

- a. Utilizando el modelo translog (3), proporcione una expresión para la elasticidad de Y con respecto a K . Calcule el valor estimado de esta elasticidad cuando se utiliza una combinación de factores productivos $(K, L) = (318, 1712)$.
- b. La especificación Cobb-Douglas (4) está anidada en la especificación translog (3) para ciertos valores de los parámetros. Contraste si la especificación Cobb-Douglas en (4) es correcta. Exponga la hipótesis nula de correcta especificación y su alternativa, proponga un estadístico para el contraste, la regla de decisión y la conclusión del mismo.

- c. Bajo la especificación en (4) contraste que la elasticidad de Y con respecto a L es igual a uno utilizando, por un lado, un contraste basado en un estadístico t y, por otro, un contraste basado en la diferencia de la suma de cuadrados de los residuos en el modelo sin restringir y en el modelo restringido. Comente sobre la validez de este contraste en cada uno de los dos casos cuando los errores son heterocedásticos
- d. Bajo la especificación Cobb-Douglas (4), contraste la hipótesis de rendimientos constantes a escala.

Modelo 1: OLS, using observations 1–27

Dependent variable: $\log Y$

	Coefficient	Std. Error	t -ratio	p-value
const	0.944197	2.91075	0.3244	0.7489
$\log L$	3.61364	1.54807	2.3343	0.0296
$\log K$	-1.89311	1.01626	-1.8628	0.0765
$(\frac{1}{2} \log^2 L)$	-0.964052	0.707385	-1.3628	0.1874
$(\frac{1}{2} \log^2 K)$	0.0852947	0.292609	0.2915	0.7735
$(\log K \cdot \log L)$	0.312387	0.438927	0.7117	0.4845
Mean dependent var	7.443631	S.D. dependent var		0.761153
Sum squared resid	0.679927	S.E. of regression		0.179937
R^2	0.954862	Adjusted R^2		0.944114
$F(5, 21)$	88.84734	P-value(F)		2.12e-13

Modelo 2: OLS, using observations 1–27

Dependent variable: $\log Y$

	Coefficient	Std. Error	t -ratio	p-value
const	1.17064	0.326782	3.5823	0.0015
$\log L$	0.602999	0.125954	4.7875	0.0001
$\log K$	0.375710	0.0853459	4.4022	0.0002
Mean dependent var	7.443631	S.D. dependent var		0.761153
Sum squared resid	0.851634	S.E. of regression		0.188374
R^2	0.943463	Adjusted R^2		0.938751
$F(2, 24)$	200.2489	P-value(F)		1.07e-15

Modelo 3: OLS, using observations 1–27

Dependent variable: $\log(Y/K)$

	Coefficient	Std. Error	t -ratio	p-value
const	1.17064	0.326782	3.5823	0.0015
$\log L$	0.602999	0.125954	4.7875	0.0001
$\log K$	-0.624290	0.0853459	-7.3148	0.0000
Mean dependent var	-0.002291	S.D. dependent var		0.356159
Sum squared resid	0.851634	S.E. of regression		0.188374
R^2	0.741779	Adjusted R^2		0.720260
$F(2, 24)$	34.47179	P-value(F)		8.79e-08

Modelo 4: OLS, using observations 1–27

Dependent variable: $\log(Y/K)$

	Coefficient	Std. Error	<i>t</i> -ratio	p-value
const	1.17064	0.326782	3.5823	0.0015
$\log(L/K)$	0.602999	0.125954	4.7875	0.0001
$\log K$	-0.000902242	0.00392627	-0.2298	0.8202
Mean dependent var	-0.002291	S.D. dependent var		0.356159
Sum squared resid	0.851634	S.E. of regression		0.188374
R^2	0.741779	Adjusted R^2		0.720260
$F(2, 24)$	34.47179	P-value(F)		8.79e-08

Modelo 5: OLS, using observations 1–27

Dependent variable: $\log(Y/L)$

	Coefficient	Std. Error	<i>t</i> -ratio	p-value
const	0.674270	0.333598	2.0212	0.0541
$\log K$	0.135068	0.044442	3.0392	0.0055
Mean dependent var	1.679979	S.D. dependent var		0.251845
Sum squared resid	1.204167	S.E. of regression		0.219469
R^2	0.269790	Adjusted R^2		0.240581
$F(1, 25)$	9.236706	P-value(F)		0.005494

3. En un estudio sobre los determinantes de los salarios de las mujeres se considera el siguiente modelo

$$\begin{aligned} \log wage = & \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 exper^2 + \beta_4 kidslt6 + \beta_5 kidsge6 + \beta_6 nwifeinc \\ & + \beta_7 nkidslt6 + \beta_8 nkidsge6 + \beta_9 (nwifeinc \cdot kidslt6) + \beta_{10} (nwifeinc \cdot kidsge6) + U, \end{aligned}$$

donde *wage* es el salario hora estimado desde los ingresos totales, *educ* son años de educación, *exper* los años de experiencia laboral, *kidslt6* una variable binaria que toma el valor uno si tiene algún hijo menor de 6 años y cero en otro caso, *kidsge6* es una binaria igual a uno si tiene algún hijo entre 6 y 18 años, *nkidslt6* es el número de hijos menores de 6 años, *nkidsge6* es el número de hijos entre 6 y 18 años, y *nwifeinc* son el total de ingresos mensuales del hogar en dólares exceptuando lo aportado por la mujer. Como se espera que U esté correlacionado con *educ*, se utiliza como instrumento los años de educación de la madre (*meduc*). Utilizando los datos de Mroz (1987, *Econometrica* 55, 765-799) se ha estimado el modelo por mínimos cuadrados en dos etapas, utilizando como variable instrumental *meduc*. También reportamos los resultados del ajuste MCO de la forma reducida de *educ*. Suponemos que $Cov(meduc, U) = 0$ en lo que sigue.

- Calcule la variación porcentual estimada del salario (*wage*) cuando la experiencia laboral (*exper*) aumenta en un uno por ciento para mujeres con 10 años de experiencia laboral permaneciendo el resto de variables explicativas constantes. Calcule la elasticidad del salario estimada con respecto a *nwifeinc* para mujeres con un hijo de 5 años y otro de 10, un valor de *nwifeinc* de 200 dólares al mes y una experiencia laboral (*exper*) de 6 años.
- Contraste si el instrumento utilizado es relevante. Suponga que queremos hacer que el rendimiento de la educación dependa del número de hijos incluyendo los efectos interactivos $educ \cdot nkidslt6$ y $educ \cdot nkidsge6$; ¿qué instrumentos puede utilizar para estimar el modelo con estas dos nuevas variables explicativas?
- Suponga que tenemos información sobre los años de educación del padre (*feduc*), y sabemos que $Cov(feduc, U) = 0$. ¿Cómo contrastaría que los instrumentos *meduc* y *feduc* son exógenos?

Describa la hipótesis nula y la alternativa, el estadístico de contraste, la distribución bajo la hipótesis nula, la regla de decisión y cómo lo llevaría a cabo.

Modelo 6: TSLS, using observations 1–753

Dependent variable: *lwage*

Instrumented: educ. Instruments: *meduc*

	Coefficient	Std. Error	<i>t</i> -ratio	p-value
<i>const</i>	0.0105551	0.0047843	2.2062	0.0136
<i>educ</i>	0.0440428	0.0220666	1.9959	0.0229
<i>exper</i>	0.0437386	0.0139363	3.1385	0.0017
<i>exper</i> ²	−0.000845089	0.000411675	−2.0528	0.0401
<i>kidslt6</i>	0.0160184	0.283771	0.0564	0.9550
<i>kidsge6</i>	−0.0333846	0.0406833	−0.8206	0.4119
<i>nwifeinc</i>	0.0143522	0.00605226	2.3714	0.0177
<i>nkidslt6</i>	−0.0128727	0.386999	−0.0333	0.9735
<i>nkidsge6</i>	0.193369	0.173542	1.1143	0.2652
<i>nwifeinc</i> × <i>kidslt6</i>	0.000744505	0.00872943	0.0853	0.9320
<i>nwifeinc</i> × <i>kidsge6</i>	−0.00835355	0.00693235	−1.2050	0.2282

Modelo 7: OLS, using observations 1–753

Dependent variable: *educ*

	Coefficient	Std. Error	<i>t</i> -ratio	p-value
<i>const</i>	7.92157	0.354448	22.3490	0.0000
<i>exper</i>	0.0993556	0.0258172	3.8484	0.0001
<i>exper</i> ²	−0.00201766	0.000838280	−2.4069	0.0163
<i>nkidslt6</i>	0.582198	0.354473	1.6424	0.1009
<i>nkidsge6</i>	−0.0508087	0.0812421	−0.6254	0.5319
<i>nwifeinc</i>	0.0593655	0.0115765	5.1281	0.0000
<i>kidslt6</i>	−0.0626910	0.540865	−0.1159	0.9078
<i>kidsge6</i>	0.114852	0.369433	0.3109	0.7560
<i>nwifeinc</i> × <i>kidslt6</i>	−0.0106848	0.0146770	−0.7280	0.4668
<i>nwifeinc</i> × <i>kidsge6</i>	−0.0111834	0.0142899	−0.7826	0.4341
<i>meduc</i>	0.274200	0.0216502	12.6650	0.0000
Mean dependent var	12.28685	S.D. dependent var	2.280246	
Sum squared resid	2827.628	S.E. of regression	1.952132	
<i>R</i> ²	0.276829	Adjusted <i>R</i> ²	0.267083	
<i>F</i> (10, 742)	28.40364	P-value(<i>F</i>)	2.94e−46	

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