

*Econometría 06216*  
*Examen Parcial #2*  
*Cali, Sábado 2 de Abril de 2011*

**Profesores: Julio César Alonso C. - Carlos Giovanni González E.**

Estudiante: \_\_\_\_\_  
 Código: \_\_\_\_\_

**Instrucciones:**

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de 9 páginas; además, deben tener 2 páginas de fórmulas.
3. El examen consta de 4 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas está expresado al lado de cada pregunta.
4. Escriba su respuesta en las hojas suministradas, marque cada una de las hojas con su nombre. NO responda en las hojas de preguntas. No se dará crédito por respuestas consignadas en las hojas de preguntas.
5. El examen está diseñado para dos horas, pero ustedes tienen 3 horas para trabajar en él.
6. Recuerde que no se tolerará ningún tipo de deshonestidad académica. En especial usted no puede emplear ningún tipo de ayuda diferente a la que se le entrega con este examen.
7. El uso de calculadoras está prohibido.
8. No se aceptarán reclamos de exámenes no escritos a lapicero (escritos a lápiz).
9. Al finalizar su examen entregue sus hojas de respuesta, así como las hojas de preguntas.
10. ¡Asigne su tiempo de forma eficiente!

Suerte.

**1 Falso o Verdadero (25 puntos en total, 5 puntos cada subparte)**

Diga si las siguientes afirmaciones son falsas o verdaderas y explique en dos o tres líneas su respuesta. (No se dará ningún crédito a respuestas sin justificación.)

- a) Un estudiante, preparándose para escribir su trabajo de grado, encuentra una nueva teoría para explicar el comportamiento de la demanda de energía de los hogares. El modelo corresponde a:  $y = X\beta + \varepsilon$ . En éste,  $X$  está conformado por una columna de "unos", además, por una columna que contiene una dummy que toma el valor de 1 cuando el jefe de hogar es una mujer y cero en caso contrario, y las otras columnas corresponden a las variables que la teoría económica establece como determinantes de la demanda de un bien. Por otro lado, este modelo quiere ser empleado para explicar el comportamiento de la demanda en una población remota que esta conformada por hogares donde la cabeza de hogar son todas mujeres. Dada esta situación, el estudiante afirma: "si se eliminase la dummy, posiblemente se estaría incurriendo en un problema econométrico de sesgo por omisión de variables omitidas, por eso el modelo debe correrse tal como está expresado arriba.". ¿Es la afirmación entre comillas verdadera o falsa?
- b) En presencia de heteroscedasticidad y una muestra pequeña, no se debe emplear la corrección de White, pues no podemos garantizar que los errores sigan una distribución normal, ni podemos emplear el teorema del límite central. ¿Es esta afirmación correcta?
- c) La multicolinealidad (no perfecta) es un problema de grado. Entre mayor sea el nivel de colinealidad entre las variables independientes, mayor es la probabilidad de obtener estimadores MCO que no son MELI.
- d) Sobre los problemas de auto-correlación y de heteroscedasticidad se puede afirmar: "En la práctica, el estimador de Mínimos Cuadrados Generalizados siempre se emplea pues: i) es el estimador MELI de los coeficientes de un modelo de regresión y ii) permite chequear si los residuos de este modelo cumplen o no los supuestos del teorema de Gauss-Markov".
- e) Cuando la muestra es lo suficientemente grande, el teorema del límite central garantiza que los estimadores MCO sigan una distribución normal, y se garantiza que los residuos estimados sigan una distribución normal.

**2 Selección Múltiple (15 puntos en total, 5 puntos cada subparte)**

Determine cuál de las siguientes respuestas es la correcta. Escoja la mejor opción y explique en dos o tres líneas su respuesta. (No se dará ningún crédito a respuestas sin justificación. Consigne su respuesta en la hoja de respuestas suministrada.)

- a) Un estudiante de administración estima un modelo econométrico para su proyecto de grado, pero no sabe si su regresión es concluyente o no. Por lo que decide consultarle a un economista. Este después de escuchar al administrador le aconseja: “Deberías tener en cuenta que los problemas que afectan la inferencia estadística del modelo también afectan al estadístico F-global”. Por lo que el administrador le pregunta al economista ¿Cuáles son algunos de los problemas econométricos que siempre afectan el F-Global en una regresión por MCO?
  - a. Omisión de variables explicativas relevantes.
  - b. Multicolinealidad, heteroscedasticidad y autocorrelación.
  - c. Muestras pequeñas, heteroscedasticidad y autocorrelación.
  - d. Error de medición en X, heteroscedasticidad y autocorrelación.
  - e. Ninguna de las anteriores.
- b) Con información recogida anualmente, durante las últimas dos décadas, se ha obtenido los siguientes resultados sobre la inversión en vivienda:

$$\ln \hat{Y}_t = -0,913 - 0,381 \ln P_t + 0,0098t; R^2 = 0,341; DW = 1,54$$

donde,  $Y_t$ : Inversión en vivienda;  $P_t$ : Índice de precios de vivienda;  $t: 0, 1, 2, \dots, 20$ .

Los datos son una serie de tiempo y se cree que posiblemente el modelo tenga problemas de autocorrelación. Entonces, realizamos la prueba de Durbin Watson y tenemos:

$\hat{\rho}_t = \rho_{t-1} + \varepsilon_t$ ; Y la prueba de hipótesis sería:  $H_0: \rho = 0$  vs  $H_a: \rho > 0$  (Nota: para  $K=2$  y  $T=21$  se tiene que  $d_L = 0,863$ ;  $d_U = 1,277$ .) Por lo tanto, se puede afirmar que:

- a. Existe autocorrelación positiva.
  - b. Existe autocorrelación negativa.
  - c. No existe autocorrelación.
  - d. El DW no es concluyente.
  - e. Ninguna de las anteriores.
- c) Un estudiante de economía se enfrenta a un taller en el que debe estimar un modelo econométrico que se caracteriza por la existencia de variables independientes omitidas correlacionadas con las incluidas y la presencia de un error de medición en la variable X1. El estudiante está seguro que si estima ese modelo, obtendrá estimadores MCO que son:
- a. Insegados, consistentes e ineficientes.
  - b. Insegados, consistentes y eficientes.

- c. Sesgados e inconsistentes.
- d. Sesgados y consistentes.
- e. Ninguna de las anteriores.

**3. (30 puntos)**

El jefe de estudios económicos de una multinacional productora de video juegos para teléfonos celulares desea estimar la función de demanda del mercado. Para tal fin, emplea las siguientes variables:  $\ln(Q_i)$  corresponde al logaritmo del número de video juegos que compra el usuario  $i$ ;  $P_i$  es el precio del último video juego comprado por el usuario  $i$  (medido en dólares);  $I_i$  representa el ingreso de cada usuario (medido en miles de dólares) y la variable  $D_i$  es una variable dummy que toma el valor de 1 si el usuario  $i$  es hombre y cero en caso contrario. Con la información disponible y con los conocimientos sobre econometría, el jefe de estudios económicos decide estimar tres modelos que se reportan al final en las tablas: Tabla 1, Tabla 2 y Tabla 3.

a) Escriba el modelo estimado por el jefe de estudios económicos, presentado en la Tabla 2. (3 puntos)

b) Teniendo en cuenta el enunciado y suponiendo que el modelo estimado en la Tabla 2 corresponde a un modelo teórico. ¿Existe algún problema econométrico en el modelo estimado y reportado en la Tabla 2? En caso de que exista uno o varios problemas menciónelo y realice las pruebas necesarias. (Sea lo más preciso y emplee únicamente la información disponible). (7 puntos)

c) Teniendo en cuenta toda la información disponible. ¿Cómo corregiría el problema anterior, si es que existía uno? Si existe un problema resuélvalo con el menor número de regresiones o cálculos posibles. Si no existe ningún problema explique las bondades del modelo. Sea lo más claro posible y en caso de poder hacer pruebas efectúelas. (7 puntos)

d) Compare los tres modelos estimados (Tabla 1, 2 y 3) por el jefe de estudios económicos de la multinacional y elija el mejor. Explique su decisión. (6 puntos)

e) Teniendo en cuenta los resultados de los apartados anteriores. Interprete los coeficientes del mejor modelo y tenga en cuenta su significancia. (7 puntos)

**4. (30 Puntos)**

El departamento de mercadeo de una firma productora de cuadernos ha encontrado que la mejor manera de explicar el comportamiento del logaritmo de la cantidad vendida  $y_t$  (en la firma acostumbran medir las cantidades de cuadernos en miles de cajas de 100 unidades) de su producto estrella es:

$$y_t = \beta_1 + \beta_2 X_{1,t} + \beta_3 X_{1,t}^2 + \beta_4 X_{2,t} + \varepsilon_t \quad t = 1, 2, \dots, 100$$

donde  $X_{2,t}$  representa el precio de esa referencia de cuadernos en el periodo t (medido en miles de pesos) y  $X_{1,t}$  denota el número de centímetros cuadrados de avisos en revistas en el periodo t. Además se sabe que:

$$E[\varepsilon_i] = 0 \quad \text{Var}[\varepsilon_i] = \sigma^2 X_{3,t} \quad E[\varepsilon_j \varepsilon_i] = 0 \quad \forall i \neq j$$

y  $X_{3,t} = 4X_{2,t}$ .

- a) Enumere claramente los supuestos que no se cumplen para poder obtener conclusiones válidas sobre los estimadores de Mínimos Cuadrados Ordinarios (MCO)? (2 puntos)
- b) El departamento de mercadeo recogió la información necesaria para estimar el modelo y obtener estimadores MELI. La información recolectada se sintetiza en la siguiente matriz  $X^T X$  y  $X^T y$ :

$$X^T X = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 200 & 0 & 0 \\ 0 & 0 & 50 & 0 \\ 0 & 0 & 0 & 20 \end{bmatrix} \quad X^T y = \begin{bmatrix} 60 \\ 100 \\ 2 \\ -20 \end{bmatrix}$$

Explique claramente a que corresponde los elementos (1,1), (2,1), (3,1) y (4,4) de la matriz  $X^T X$ . Exprese su respuesta únicamente en términos de  $X_{1,t}$  y  $X_{2,t}$  (8 puntos – 2 puntos cada uno)

- c) Encuentre los estimadores MELI de los  $\beta$ 's del modelo. Muestre claramente cuál es el valor estimado de cada uno de los parámetros del modelo. (6 Puntos)
- d) Interprete el significado de cada uno de los coeficientes estimados. (8 Puntos)
- e) (Esta pregunta es independiente de la pregunta anterior. Es decir, parta del modelo inicial). Un nuevo pasante en el departamento de mercadeo sugiere estimar los siguientes dos modelos:

Modelo A  $y_i = \beta_1 + \beta_2 X_{1,t} + \beta_3 X_{1,t}^2 + \beta_4 D1_t X_{2,t} + \beta_5 D2_t X_{2,t} + \varepsilon_i$

Modelo B  $y_i = \beta_1 + \beta_2 X_{1,t} + \beta_3 X_{1,t}^2 + \varepsilon_i$

donde  $D1_t$  es una variable dummy que toma el valor de uno después de la apertura del mercado a la competencia internacional y cero en caso contrario. Y  $D2_t$  es otra variable dummy que toma el valor de uno antes de la apertura de la apertura del mercado a la competencia internacional y cero en caso contrario.

El director de estudios afirma: “la sugerencia del pasante no tiene mucho sentido. En ambos casos los estimadores MCO de los  $\beta$ s no serán Meli, aún si se soluciona cualquier problema de autocorrelación o heteroscedasticidad que se presenten”.  
¿Está usted de acuerdo con la afirmación del director? Sea lo más claro posible y argumente por qué el método MCO produce o no estimadores Meli de los  $\beta$ s en cada modelo (el A y el B). (6 Puntos)

**Tabla 1**

Dependent variable:  
 $Y = Q/\sigma_i$

Characteristics:  
 First observation = 1  
 Last observation = 50  
 Number of usable observations: 50

X variables:  
 $X(1) = P/\sigma_i$   
 $X(2) = I/\sigma_i$   
 $X(3) = D/\sigma_i$   
 $X(4) = 1/\sigma_i$

Model:  
 $Y = b(1)X(1) + b(2)X(2) + b(3)X(3) + b(4)X(4) + U$   
 where U is the error term, satisfying  
 $E[U|X(1), X(2), X(3), X(4)] = 0$ .

OLS estimation results

Parameters	Estimate	t value [p-value]	H.C t-value [H.C. p-value]
b(1)	-0.30245	2.984 [0.00000]	1.543 [0.07524]
b(2)	0.14284	2.339 [0.05000]	2.530 [0.00000]
b(3)	0.05524	2.983 [0.03256]	1.229 [0.12544]
b(4)	10.20257	1.422 [0.07800]	2.578 [0.04835]

Effective sample size (n): 50  
 Variance of the residuals: 0.12151531  
 Standard error of the residuals (SER): 0.34859046  
 R-square: 0.8957  
 Adjusted R-square: 0.8726

Test for first-order autocorrelation:  
 Durbin-Watson test = 1.49

Jarque-Bera/Salmon-Kiefer test = 50.94518  
 Null hypothesis: The errors are normally distributed  
 Null distribution: Chi-square(2)  
 p-value = 0.00384  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

Breusch-Pagan test = 13,73412  
 Null hypothesis: The errors are homoskedastic  
 Null distribution: Chi-square(2)  
 p-value = 0.003285  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

**Tabla 2**

Dependent variable:  
 $Y = LN[Q]$

Characteristics:  
 First observation = 1  
 Last observation = 50  
 Number of usable observations: 50

X variables:  
 $X(1) = P$   
 $X(2) = I$   
 $X(3) = D$   
 $X(4) = 1$

Model:  
 $Y = b(1)X(1) + b(2)X(2) + b(3)X(3) + b(4)X(4) + U$   
 where U is the error term, satisfying  
 $E[U|X(1), X(2), X(3), X(4)] = 0$ .

OLS estimation results

Parameters	Estimate	t value [p-value]	H.C t-value [H.C. p-value]
b(1)	-0.31524	2.984 [0.00000]	2.543 [0.00000]
b(2)	0.15261	2.339 [0.05000]	2.530 [0.00000]
b(3)	0.05902	2.983 [0.03256]	1.229 [0.12544]
b(4)	10.20146	1.422 [0.07800]	2.578 [0.04835]

Effective sample size (n): 50  
 Variance of the residuals: 0.12151531  
 Standard error of the residuals (SER): 0.34859046  
 R-square: 0.5957  
 Adjusted R-square: 0.5726

Test for first-order autocorrelation:  
 Durbin-Watson test = 1.49

Jarque-Bera/Salmon-Kiefer test = 50.94518  
 Null hypothesis: The errors are normally distributed  
 Null distribution: Chi-square(2)  
 p-value = 0.00384  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

Breusch-Pagan test = 3,73412  
 Null hypothesis: The errors are homoskedastic  
 Null distribution: Chi-square(2)  
 p-value = 0.3285  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

**Tabla 3**

Dependent variable:  
 $Y = LN[Q]$

Characteristics:  
 First observation = 1  
 Last observation = 10  
 Number of usable observations: 10

X variables:  
 $X(1) = P$   
 $X(2) = I$   
 $X(3) = D$   
 $X(4) = 1$

Model:  
 $Y = b(1)X(1) + b(2)X(2) + b(3)X(3) + b(4)X(4) + U.$   
 where U is the error term, satisfying  
 $E[U|X(1), X(2), X(3), X(4)] = 0.$

OLS estimation results

Parameters	Estimate	t value [p-value]	H.C t-value [H.C. p-value]
b(1)	-0.32548	2.984 [0.00000]	2.543 [0.00000]
b(2)	0.12548	2.339 [0.05000]	2.530 [0.00000]
b(3)	0.05695	2.983 [0.03256]	1.229 [0.12544]
b(4)	10.21584	1.422 [0.07800]	2.578 [0.04835]

Effective sample size (n): 10  
 Variance of the residuals: 0.12151531  
 Standard error of the residuals (SER): 0.34859046  
 R-square: 0.7957  
 Adjusted R-square: 0.7726

Test for first-order autocorrelation:  
 Durbin-Watson test = 1.49

Jarque-Bera/Salmon-Kiefer test = 3.4518  
 Null hypothesis: The errors are normally distributed  
 Null distribution: Chi-square(2)  
 p-value = 0.2537  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

Breusch-Pagan test = 13,73412  
 Null hypothesis: The errors are homoskedastic  
 Null distribution: Chi-square(2)  
 p-value = 0.005687  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

*Econometría 06216*  
*Examen Parcial #2*  
*Respuestas Sugeridas*  
*Calí, Sábado 2 de Abril de 2011*

**Profesores: Julio César Alonso C. - Carlos Giovanni González E.**

Estudiante: \_\_\_\_\_  
 Código: \_\_\_\_\_

**Instrucciones:**

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de 9 páginas; además, deben tener 2 páginas de fórmulas.
3. El examen consta de 4 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas está expresado al lado de cada pregunta.
4. Escriba su respuesta en las hojas suministradas, marque cada una de las hojas con su nombre. NO responda en las hojas de preguntas. No se dará crédito por respuestas consignadas en las hojas de preguntas.
5. El examen está diseñado para dos horas, pero ustedes tienen 3 horas para trabajar en él.
6. Recuerde que no se tolerará ningún tipo de deshonestidad académica. En especial usted no puede emplear ningún tipo de ayuda diferente a la que se le entrega con este examen.
7. El uso de calculadoras está prohibido.
8. No se aceptarán reclamos de exámenes no escritos a lapicero (escritos a lápiz).
9. Al finalizar su examen entregue sus hojas de respuesta, así como las hojas de preguntas.
10. ¡Asigne su tiempo de forma eficiente!

Suerte.

**1 Falso o Verdadero (25 puntos en total, 5 puntos cada subparte)**

Diga si las siguientes afirmaciones son falsas o verdaderas y explique en dos o tres líneas su respuesta. (No se dará ningún crédito a respuestas sin justificación.)

- a) Un estudiante, preparándose para escribir su trabajo de grado, encuentra una nueva teoría para explicar el comportamiento de la demanda de energía de los hogares. El modelo corresponde a:  $y = X\beta + \varepsilon$ . En éste,  $X$  está conformado por una columna de “unos”, además, por una columna que contiene una dummy que toma el valor de 1 cuando el jefe de hogar es una mujer y cero en caso contrario, y las otras columnas corresponden a las variables que la teoría económica establece como determinantes de la demanda de un bien. Por otro lado, este modelo quiere ser empleado para explicar el comportamiento de la demanda en una población remota que está conformada por hogares donde la cabeza de hogar son todas mujeres. Dada esta situación, el estudiante afirma: “*si se eliminase la dummy, posiblemente se estaría incurriendo en un problema econométrico de sesgo por omisión de variables omitidas, por eso el modelo debe correrse tal como está expresado arriba.*”. ¿Es la afirmación entre comillas verdadera o falsa?

La afirmación es falsa, dado que en este caso no tiene sentido incluir la dummy, pues se genera multicolinealidad perfecta al ser toda la población mujeres cabeza de hogar.

Por eso, el modelo deberá correrse sin la variable dummy. Y en este caso el problema de omisión de variables será irrelevante.

Nota si no se menciona el problema de multicolinealidad perfecta, entonces no se dará crédito.

- b) En presencia de heteroscedasticidad y una muestra pequeña, no se debe emplear la corrección de White, pues no podemos garantizar que los errores sigan una distribución normal, ni podemos emplear el teorema del límite central. ¿Es esta afirmación correcta?

Falso, la cláusula inicial es correcta (En presencia de heteroscedasticidad y una muestra pequeña, no se debe emplear la corrección de White), pero la razón no lo es. En muestras pequeñas dicho estimador será sesgado. Por tanto la razón para no emplear el estimador White no es la normalidad, sino el sesgo que presenta en muestras pequeñas.

- c) La multicolinealidad (no perfecta) es un problema de grado. Entre mayor sea el nivel de colinealidad entre las variables independientes, mayor es la probabilidad de obtener estimadores MCO que no son MELI.

Falso. La multicolinealidad en efecto es un problema de grado, cuando no es perfecta no se viola ningún supuesto del teorema de Gauss-Markov. Por tanto los estimadores son MELI.

Es importante anotar que en presencia de multicolinealidad, la varianza de los estimadores tiende a ser grande, con respecto al caso en el que no existe multicolinealidad. Pero dado que existe multicolinealidad, no es posible encontrar otro estimador lineal insesgado que tenga una varianza más pequeña que los MCO.

- d) Sobre los problemas de auto-correlación y de heteroscedasticidad se puede afirmar: “En la práctica, el estimador de Mínimos Cuadrados Generalizados siempre se empleara pues: i) es el estimador MELI de los coeficientes de un modelo de regresión y ii) permite chequear si los residuos de este modelo cumplen o no los supuestos del teorema de Gauss-Markov”.

Falso. Si bien el estimador de MCG es el estimador MELI de los coeficientes de un modelo de regresión y permite chequear si los residuos de este modelo cumplen o no los supuestos del teorema de Gauss-Markov, no es correcto afirmar que “En la práctica, el estimador de Mínimos Cuadrados Generalizados siempre se empleara”. Recuerden que para emplear dicho estimador, será necesario conocer la naturaleza de la matriz de varianzas y covarianzas del error. Y esto normalmente no lo conocemos, por eso, si bien los MCG son una buena idea, estos son difíciles de emplear en la práctica.

- e) Cuando la muestra es lo suficientemente grande, el teorema del límite central garantiza que los estimadores MCO sigan una distribución normal, y se garantiza que los residuos estimados sigan una distribución normal.

Falso. Con el teorema del límite central se garantiza que los estimadores sigan una distribución normal, cuando la muestra es lo suficientemente grande. Pero no se puede garantizar que los errores estén distribuidos normalmente.

**2 Selección Múltiple (15 puntos en total, 5 puntos cada subparte)**

Determine cuál de las siguientes respuestas es la correcta. Escoja la mejor opción y explique en dos o tres líneas su respuesta. (No se dará ningún crédito a respuestas sin justificación. Consigne su respuesta en la hoja de respuestas suministrada.)

- a) Un estudiante de administración estima un modelo econométrico para su proyecto de grado, pero no sabe si su regresión es concluyente o no. Por lo que decide consultarle a un economista. Éste, después de escuchar al administrador, le aconseja: “Deberías tener en cuenta que los problemas que afectan la inferencia estadística sobre el modelo también afectan al estadístico F-global”. Por lo que el administrador le pregunta al economista ¿Cuáles son algunos de los problemas econométricos que siempre afectan el F-Global en una regresión por MCO?
- a. Omisión de variables explicativas relevantes.

- b. Multicolinealidad, heteroscedasticidad y autocorrelación.
- c. Muestras pequeñas, heteroscedasticidad y autocorrelación.
- d. Error de medición en X, heteroscedasticidad y autocorrelación.
- e. Ninguna de las anteriores.

Respuesta sugerida: d.

- b) Con información recogida anualmente, durante las últimas dos décadas, se ha obtenido los siguientes resultados sobre la inversión en vivienda:

$$\ln \hat{Y}_t = -0,913 - 0,381 \ln P_t + 0,0098t; R^2 = 0,341; DW=1,54$$

donde,  $Y_t$ : Inversión en vivienda;  $P_t$ : Índice de precios de vivienda;  $t: 0, 1, 2, \dots, 20$ .

Los datos son una serie de tiempo y se cree que posiblemente el modelo tenga problemas de autocorrelación. Entonces, realizamos la prueba de Durbin Watson y tenemos:  $d_t = \rho_{t-1} + \epsilon_t$ ; Y la prueba de hipótesis sería:  $H_0: \rho = 0$  vs  $H_a: \rho > 0$  (Nota: para  $K=2$  y  $T=21$  se tiene que  $d_l = 0,863$ ;  $d_u = 1,277$ .) Por lo tanto, se puede afirmar que:

- a. Existe autocorrelación positiva.
- b. Existe autocorrelación negativa.
- c. No existe autocorrelación.
- d. El DW no es concluyente.
- e. Ninguna de las anteriores.

Respuesta sugerida: c.

- c) Un estudiante de economía se enfrenta a un taller en el que debe estimar un modelo econométrico que se caracteriza por la existencia de variables independientes omitidas correlacionadas con las incluidas y la presencia de un error de medición en la variable X1. El estudiante está seguro que si estima ese modelo, obtendrá estimadores MCO que son:

- a. Insesgados, consistentes e ineficientes.
- b. Insesgados, consistentes y eficientes.
- c. Sesgados e inconsistentes.
- d. Sesgados y consistentes.
- e. Ninguna de las anteriores.

Respuesta sugerida: c.

**3. (30 puntos)**

El jefe de estudios económicos de una multinacional productora de video juegos para teléfonos celulares desea estimar la función de demanda del mercado. Para tal fin,

emplea las siguientes variables:  $\ln(Q_i)$  corresponde al logaritmo del número de video juegos que compra el usuario  $i$ ;  $P_i$  es el precio del último video juego comprado por el usuario  $i$  (medido en dólares);  $I_i$  representa el ingreso de cada usuario (medido en miles de dólares) y la variable  $D_i$  es una variable dummy que toma el valor de 1 si el usuario  $i$  es hombre y cero en caso contrario. Con la información disponible y con los conocimientos sobre econometría, el jefe de estudios económicos decide estimar tres modelos que se reportan al final en las tablas: Tabla 1, Tabla 2 y Tabla 3.

- a) Escriba el modelo estimado por el jefe de estudios económicos, presentado en la Tabla 2. (3 puntos)

$$\ln(Q_i) = \beta_0 + \beta_1 P_i + \beta_2 I_i + \beta_3 D_i + \varepsilon_i$$

$$i = 1, 2, \dots, 50.$$

- b) Teniendo en cuenta el enunciado y suponiendo que el modelo estimado en la Tabla 2 corresponde a un modelo teórico. ¿Existe algún problema econométrico en el modelo estimado y reportado en la Tabla 2? En caso de que exista uno o varios problemas menciónelos y realice las pruebas necesarias. (Sea lo más preciso y emplee únicamente la información disponible). (7 puntos)

Noten que en la estimación presentada en la Tabla 2 no existe ningún problema econométrico de los vistos en clase.

Por descartar los problemas de multicolinealidad y autocorrelación (corte transversal!!) (1 punto), por descartar los problemas de omisión de variables relevantes, errores de medición en las independientes y la distribución normal del error en muestras pequeñas (2 puntos).

Es necesario realizar la prueba de heteroscedasticidad debido a que los datos son de corte transversal. Realizar la prueba de hipótesis (1 punto), por el criterio (1 punto) y por reportar la conclusión (2 puntos).

- c) Teniendo en cuenta toda la información disponible. ¿Cómo corregiría el problema anterior, si es que existía uno? Si existe un problema resuélvalo con el menor número de regresiones o cálculos posibles. Si no existe ningún problema explique las bondades del modelo. Sea lo más claro posible y en caso de poder hacer pruebas efectúelas. (7 puntos)

Debido a que el modelo reportado en la tabla 2 no tiene problemas econométricos. Se habla de sus bondades. Mencionar que el tamaño de la muestra es lo suficientemente grande (1,5 punto), mencionar que la estimación tiene intercepto (1,5 punto). Interpretar el

$R^2$ . El 57 por ciento de las variaciones de la variable dependiente están explicadas por el modelo de regresión (5 puntos).

Noten que en la tabla 1 se presenta la estimación del modelo de la tabla 2 pero ponderado (MCP), pero el modelo de la tabla 1 no tiene la misma variable dependiente. No se dará crédito por esta respuesta.

- d) Compare los tres modelos estimados (Tabla 1, 2 y 3) por el jefe de estudios económicos de la multinacional y elija el mejor. Explique su decisión. (6 puntos)

El mejor modelo es el estimado en la Tabla 2.

El modelo de la Tabla 1 que es una ponderación del modelo, no tiene problemas econométricos, pero NO usa las variables elegidas por el jefe de estudios y no es comparable. Por lo tanto, sólo son comparables los modelos de las tablas 2 y 3 (1 punto).

Sabemos que el modelo de la tabla 2 no tiene problemas. Por otro lado, noten que el modelo de la tabla 3 tiene una muestra pequeña por lo que se debe probar la distribución normal de los errores. En este caso según la prueba de JB se distribuyen normalmente por lo que si podemos hacer inferencia. Este modelo también presenta un problema de heteroscedasticidad (1 punto, por el análisis del modelo reportado en la tabla 3).

Para elegir entre los dos modelos, se debe tener en cuenta que en el modelo reportado en la tabla 3 no se puede usar la corrección de White (2 punto). Por lo que el mejor modelo sería el de la tabla 2 (2 puntos)

- e) Teniendo en cuenta los resultados de los apartados anteriores. Interprete los coeficientes del mejor modelo y tenga en cuenta su significancia. (7 puntos)

Los coeficientes que se deben interpretar son los del modelo de la tabla 2. Si se interpreta el modelo reportado en la tabla 3 no se penalizará. Pero si se reporta el de la tabla 1 no se le dará crédito.

$\hat{\beta}_0 = 10.20146$  Significativo al 95%. No tiene interpretación económica.

$\hat{\beta}_1 = -0.31524$  Significativo al 99%. Ante un aumento de un dólar en el precio del último juego comprado por el usuario, se espera una caída en la demanda de juegos del 31.52 por ciento.

$\hat{\beta}_2 = 0.15261$  Significativo al 90%.

Ante un aumento de mil dólares en el ingreso del usuario, se espera un aumento en la demanda de juegos del 15.26 por ciento.

Con un 95 o 99% de confianza, el ingreso del usuario no tiene efecto en la demanda de video juegos.

$\hat{\beta}_3 = 0.05902$  No es significativo con un 99% de confianza. El género del usuario (ser hombre) no tiene influencia sobre la demanda de video juegos para teléfonos celulares.

**4. (30 Puntos)**

El departamento de mercadeo de una firma productora de cuadernos ha encontrado que la mejor manera de explicar el comportamiento del logaritmo de la cantidad vendida  $y_t$  (en la firma acostumbran medir las cantidades de cuadernos en miles de cajas de 100 unidades) de su producto estrella es:

$$y_t = \beta_1 + \beta_2 X_{1,t} + \beta_3 X_{1,t}^2 + \beta_4 X_{2,t} + \varepsilon_t \quad t = 1, 2, \dots, 100$$

donde  $X_{2,t}$  representa el precio de esa referencia de cuadernos en el periodo t (medido en miles de pesos) y  $X_{1,t}$  denota el número de centímetros cuadrados de avisos en revistas en el periodo t. Además se sabe que:

$$E[\varepsilon_t] = 0 \quad \text{Var}[\varepsilon_t] = \sigma^2 X_{3,t} \quad E[\varepsilon_i \varepsilon_j] = 0 \quad \forall i \neq j$$

y  $X_{3,t} = 4X_{2,t}$ .

- a) Enumere claramente los supuestos que no se cumplen para poder obtener conclusiones válidas sobre los estimadores de Mínimos Cuadrados Ordinarios (MCO)? (2 puntos)

*Es importante anotar que para obtener conclusiones se necesita que los estimadores MCO sean MELI y que estos sigan una distribución normal para poder hacer inferencia sobre ellos. En ese orden de ideas, dado que la muestra es lo suficientemente grande, únicamente no se cumple el supuesto de homoscedasticidad. (2 puntos)*

*Se restará un punto si se mencionan otros problemas que no existen.*

- b) El departamento de mercadeo recogió la información necesaria para estimar el modelo y obtener estimadores MELI. La información recolectada se sintetiza en la siguiente matriz  $X^T X$  y  $X^T y$ :

$$X^T X = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 200 & 0 & 0 \\ 0 & 0 & 50 & 0 \\ 0 & 0 & 0 & 20 \end{bmatrix} \quad X^T y = \begin{bmatrix} 60 \\ 100 \\ 2 \\ -20 \end{bmatrix}$$

Explique claramente a que corresponde los elementos (1,1), (2,1), (3,1) y (4,4) de la matriz  $X^T X$ . Expresé su respuesta únicamente en términos de  $X_{1,t}$  y  $X_{2,t}$  (8 puntos – 2 puntos cada uno)

*En este caso se viola el supuesto de homoscedasticidad. Es decir el término de error no tiene varianza constante. El problema se puede solucionar fácilmente empleando los mínimos cuadrados ponderados. Es decir, dividiendo todo el modelo por  $\sqrt{X_{2,t}}$ .*

$$\frac{y_t}{\sqrt{X_{2,t}}} = \frac{\beta_1}{\sqrt{X_{2,t}}} + \beta_2 \frac{X_{1,t}}{\sqrt{X_{2,t}}} + \beta_3 \frac{X_{1,t}^2}{\sqrt{X_{2,t}}} + \beta_4 \frac{X_{2,t}}{\sqrt{X_{2,t}}} + \frac{\varepsilon_t}{\sqrt{X_{2,t}}}$$

$$\frac{y_t}{\sqrt{X_{2,t}}} = \frac{\beta_1}{\sqrt{X_{2,t}}} + \beta_2 \frac{X_{1,t}}{\sqrt{X_{2,t}}} + \beta_3 \frac{X_{1,t}^2}{\sqrt{X_{2,t}}} + \beta_4 \sqrt{X_{2,t}} + \frac{\varepsilon_t}{\sqrt{X_{2,t}}}$$

*Así tendremos que*

$$\text{Var} \left[ \frac{\varepsilon_t}{\sqrt{X_{2,t}}} \right] = \frac{1}{X_{2,t}} \text{Var}[\varepsilon_t] = 4\sigma^2$$

*Y por tanto el problema de heteroscedasticidad ha sido solucionado.*

*El modelo entonces será:*

$$\frac{y_t}{\sqrt{X_{2,t}}} = \frac{\beta_1}{\sqrt{X_{2,t}}} + \beta_2 \frac{X_{1,t}}{\sqrt{X_{2,t}}} + \beta_3 \frac{X_{1,t}^2}{\sqrt{X_{2,t}}} + \beta_4 \sqrt{X_{2,t}} + \frac{\varepsilon_t}{\sqrt{X_{2,t}}}$$

*En este caso tenemos:*

*Elemento (1,1):*  $\sum_{i=1}^{100} \frac{1}{X_{2,t}}$

*Elemento (2,1):*  $\sum_{i=1}^{100} \frac{X_{1,t}}{X_{2,t}}$

*Elemento (3,1):*  $\sum_{i=1}^{100} \frac{(X_{1,t})^2}{X_{2,t}}$

*Elemento (4,4):*  $\sum_{i=1}^{100} X_{2,t}$

- c) Encuentre los estimadores MELI de los  $B$ 's del modelo. Muestre claramente cuál es el valor estimado de cada uno de los parámetros del modelo. (6 Puntos)

*En este caso tenemos que:*

$$\beta_{\text{hat}} = (X^T X)^{-1} \cdot X^T y$$

$$\begin{bmatrix} 1/100 & 0 & 0 & 0 \\ 0 & 1/200 & 0 & 0 \\ 0 & 0 & 1/50 & 0 \\ 0 & 0 & 0 & 1/20 \end{bmatrix} \begin{bmatrix} 60 \\ 100 \\ 2 \\ -20 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 1/2 \\ 1/25 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 1/2 \\ 1/25 \\ -1 \end{bmatrix}$$

d) Interprete el significado de cada uno de los coeficientes estimados. (8 Puntos)

$\hat{\beta}_1 = 3/5$  no tiene interpretación económica. (2 puntos)

Noten que

$$\frac{\partial Q}{\partial X_{1,t}} = Q(\beta_2 + 2\beta_3 X_{1,t})$$

$$\frac{\Delta\%Q}{\Delta X_{1,t}} = 100(\beta_2 + 2\beta_3 X_{1,t})$$

Por tanto

$\hat{\beta}_2 = 1/2$  Un aumento de un centímetro cuadrado en los avisos en revistas en el periodo t tendrá un efecto de 50 por ciento más en la demanda de cuadernos independiente del número de centímetros cuadrados de avisos que se tenga. En otras palabras, la parte constante del efecto porcentual de un centímetro más de avisos es de 50 puntos porcentuales. (2 puntos)

$\hat{\beta}_3 = 1/25$  Un aumento de un centímetro cuadrado en los avisos en revistas en el periodo t aumentará el efecto marginal de un centímetro cuadrado en el cambio porcentual de la demanda (medido en puntos porcentuales) en 8 puntos porcentuales (100\*2\*1/25). (2 puntos)

$\hat{\beta}_4 = -1$ . Un aumento de mil pesos en el precio de esa referencia de cuadernos en el periodo t (medido en miles de pesos) implica una caída de la demanda del 100%

e) (Esta pregunta es independiente de la pregunta anterior. Es decir, parta del modelo inicial). Un nuevo pasante en el departamento de mercadeo sugiere estimar los siguientes dos modelos:

Modelo A  $y_t = \beta_1 + \beta_2 X_{1,t} + \beta_3 X_{1,t}^2 + \beta_4 D1_t X_{2,t} + \beta_5 D2_t X_{2,t} + \varepsilon_t$

Modelo B  $y_t = \beta_1 + \beta_2 X_{1,t} + \beta_3 X_{1,t}^2 + \varepsilon_t$

donde  $D1_t$  es una variable dummy que toma el valor de uno después de la apertura del mercado a la competencia internacional y cero en caso contrario. Y  $D2_t$  es otra variable dummy que toma el valor de uno antes de la apertura ~~de la apertura~~ del mercado a la competencia internacional y cero en caso contrario.

El director de estudios afirma: “la sugerencia del pasante no tiene mucho sentido. En ambos casos los estimadores MCO de los  $\beta$ s no serán Meli, aún si se soluciona cualquier problema de autocorrelación o heteroscedasticidad que se presenten”.

¿Está usted de acuerdo con la afirmación del director? Sea lo más claro posible y argumente por qué el método MCO produce o no estimadores Meli de los  $\beta$ s en cada modelo (el A y el B). (6 Puntos)

Noten que el modelo A no presenta problemas de multicolinealidad y la única diferencia con el modelo original es incluir la presencia de una variable dummy. Así, no existe ninguna razón por la cual se pueda pensar que los estimadores MCO de los  $\beta$ s no serán Meli, aún si se soluciona cualquier problema de autocorrelación o heteroscedasticidad que se presenten. (2.5 Puntos)

Por otro lado, el modelo B implica omitir una variable explicatoria. Si  $X_{2,t}$  no presenta una correlación con  $X_{1,t}$  o con  $X_{1,t}^2$ , entonces los estimadores MCO (o GLS) no tendrán problemas de sesgo de omisión. En caso contrario existirá un sesgo. (2.5 Puntos)

Por lo tanto la afirmación del director no es correcta. (1 Punto)

**Tabla 1**

Dependent variable:  
 $Y = Q/\sigma_t$

Characteristics:  
 First observation = 1  
 Last observation = 50  
 Number of usable observations: 50

X variables:  
 $X(1) = P/\sigma_t$   
 $X(2) = I/\sigma_t$   
 $X(3) = D/\sigma_t$   
 $X(4) = 1/\sigma_t$

Model:  
 $Y = b(1)X(1) + b(2)X(2) + b(3)X(3) + b(4)X(4) + U$   
 where U is the error term, satisfying  
 $E[U|X(1), X(2), X(3), X(4)] = 0$ .

OLS estimation results

Parameters	Estimate	t value [p-value]	H.C t-value [H.C. p-value]
b(1)	-0.30245	2.984 [0.00000]	1.543 [0.07524]
b(2)	0.14284	2.339 [0.05000]	2.530 [0.00000]
b(3)	0.05524	2.983 [0.03256]	1.229 [0.12544]
b(4)	10.20257	1.422 [0.07800]	2.578 [0.04835]

Effective sample size (n): 50  
 Variance of the residuals: 0.12151531  
 Standard error of the residuals (SER): 0.34859046  
 R-square: 0.8957  
 Adjusted R-square: 0.8726

Test for first-order autocorrelation:  
 Durbin-Watson test = 1.49

Jarque-Bera/Salmon-Kiefer test = 50.94518  
 Null hypothesis: The errors are normally distributed  
 Null distribution: Chi-square(2)  
 p-value = 0.00384  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

Breusch-Pagan test = 13,73412  
 Null hypothesis: The errors are homoskedastic  
 Null distribution: Chi-square(2)  
 p-value = 0.003285  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

**Tabla 2**

Dependent variable:  
 $Y = LN[Q]$

Characteristics:  
 First observation = 1  
 Last observation = 50  
 Number of usable observations: 50

X variables:  
 $X(1) = P$   
 $X(2) = I$   
 $X(3) = D$   
 $X(4) = 1$

Model:  
 $Y = b(1)X(1) + b(2)X(2) + b(3)X(3) + b(4)X(4) + U$   
 where U is the error term, satisfying  
 $E[U|X(1), X(2), X(3), X(4)] = 0$ .

OLS estimation results

Parameters	Estimate	t value [p-value]	H.C t-value [H.C. p-value]
b(1)	-0.31524	2.984 [0.00000]	2.543 [0.00000]
b(2)	0.15261	2.339 [0.05000]	2.530 [0.00000]
b(3)	0.05902	2.983 [0.03256]	1.229 [0.12544]
b(4)	10.20146	1.422 [0.07800]	2.578 [0.04835]

Effective sample size (n): 50  
 Variance of the residuals: 0.12151531  
 Standard error of the residuals (SER): 0.34859046  
 R-square: 0.5957  
 Adjusted R-square: 0.5726

Test for first-order autocorrelation:  
 Durbin-Watson test = 1.49

Jarque-Bera/Salmon-Kiefer test = 50.94518  
 Null hypothesis: The errors are normally distributed  
 Null distribution: Chi-square(2)  
 p-value = 0.00384  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

Breusch-Pagan test = 3,73412  
 Null hypothesis: The errors are homoskedastic  
 Null distribution: Chi-square(2)  
 p-value = 0.3285  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

**Tabla 3**

Dependent variable:  
 $Y = LN[Q]$

Characteristics:  
 First observation = 1  
 Last observation = 10  
 Number of usable observations: 10

X variables:  
 $X(1) = P$   
 $X(2) = I$   
 $X(3) = D$   
 $X(4) = 1$

Model:  
 $Y = b(1)X(1) + b(2)X(2) + b(3)X(3) + b(4)X(4) + U.$   
 where U is the error term, satisfying  
 $E[U|X(1), X(2), X(3), X(4)] = 0.$

OLS estimation results

Parameters	Estimate	t value [p-value]	H.C t-value [H.C. p-value]
b(1)	-0.32548	2.984 [0.00000]	2.543 [0.00000]
b(2)	0.12548	2.339 [0.05000]	2.530 [0.00000]
b(3)	0.05695	2.983 [0.03256]	1.229 [0.12544]
b(4)	10.21584	1.422 [0.07800]	2.578 [0.04835]

Effective sample size (n): 10  
 Variance of the residuals: 0.12151531  
 Standard error of the residuals (SER): 0.34859046  
 R-square: 0.7957  
 Adjusted R-square: 0.7726

Test for first-order autocorrelation:  
 Durbin-Watson test = 1.49

Jarque-Bera/Salmon-Kiefer test = 3.4518  
 Null hypothesis: The errors are normally distributed  
 Null distribution: Chi-square(2)  
 p-value = 0.2537  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99

Breusch-Pagan test = 13,73412  
 Null hypothesis: The errors are homoskedastic  
 Null distribution: Chi-square(2)  
 p-value = 0.005687  
 Significance levels: 10% 5%  
 Critical values: 4.61 5.99