

Econometría 06216
Examen Final
Cali, Miércoles 17 de noviembre de 2010

Profesores: Julio César Alonso --- Carlos Giovanni González

Estudiante: _____

Código: _____

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **12** páginas; además, deben tener 2 páginas de fórmulas.
3. El examen consta de 3 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas esta 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.
5. El examen está diseñado para dos horas, pero ustedes tienen 4 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 respuestas escritas en 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!

I. Selección Múltiple (50 puntos en total, 1 punto por cada subparte)

Seleccione la opción más indicada en la hoja de respuestas que encontrará al final de este examen. Sólo se considerarán respuestas que sean consignadas en la hoja de respuestas. (No es necesario justificar su respuesta)

1. Suppose the following model $y_i = \beta_1 + \beta_2 D_i x_{1i} + \beta_3 W_i x_{1i} + \varepsilon_i$ where D_i and W_i are dummy variables, and

$$D_i = \begin{cases} 1 & \text{female} \\ 0 & \text{o.w} \end{cases} \text{ and } W_i = \begin{cases} 1 & \text{male} \\ 0 & \text{o.w} \end{cases},$$

therefore the β 's estimates will most likely be:

- a. Inconsistent.
 - b. Perfectly multicollinear.
 - c. Unbiased.
 - d. Biased.
 - e. Efficient.
2. Un econométra está convencido de la necesidad de una de los siguientes supuestos para que el estimador MCO de β sea insesgado. ¿Cuál es ese supuesto?:
- a) Que la varianza del término de error sea constante a lo largo de la muestra
 - b) Que la distribución de los errores sea normal
 - c) Que los parámetros del modelo no sean constantes a lo largo de la muestra
 - d) Que la covarianza entre cualquier par de errores sea igual a cero
 - e) Ninguno de los anteriores
3. One tailed tests are *sometimes* used to test hypotheses about regression coefficients. In which of the following circumstances?
- a. when the estimated coefficient is close to zero

- b. when you wish to use a larger significance level than 5%
- c. when the sample size is large enough to use the normal approximation to the t distribution
- d. when the estimated coefficient has the opposite sign to that predicted by theory
- e. when you are testing a hypothesis other than that the parameter equals zero

4. Considere el siguiente modelo de regresión $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$; donde, μ cumple los supuestos habituales. Y con respecto a las variables originales Y , X_1 , X_2 , el modelo es:

- a. Lineal en parámetros y en variables
 - b. No lineal en parámetros y lineal en variables
 - c. Lineal en parámetros y no lineal en variables
 - d. No lineal en parámetros y no lineal en variables
 - e. Ninguna de las anteriores
5. The fact that least squares estimator of the slope coefficients is unbiased means that:
- a. the estimated slope coefficient will always be equal to the true parameter value
 - b. the estimated slope coefficient will get closer to the true parameter value as the size of the sample increases
 - c. the estimated slope coefficient will be equal to the true parameter if the sample is large
 - d. the mean of the sampling distribution of the slope parameter is zero
 - e. if repeated samples of the same size are taken, on average their value will be equal to the true parameter
6. Un econometrista experimentado afirma que en uno de los siguientes modelos los parámetros β_1 y β_2 no podrían ser estimados por MCO (donde μ es el término de error). ¿Cuál?

- a) $Y_i = \beta_1 + \beta_2(\ln X_i)^2 + \mu_i$
- b) $Y_i = e^{(\beta_1 + \beta_2 X_i + \mu_i)}$
- c) $Y_i = \sqrt{\beta_1 + \beta_2 X_i} + \mu_i$
- d) $Y_i = e^{\beta_1} X_i^{\beta_2} e^{\mu_i}$
- e) Ninguna de las anteriores

7. Consider the classical econometrics model with noexplanatory variables ($n > 1$), $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$. Which of the following statements is false?
- a. The sample average is a linear estimator.
 - b. The sample average and the OLS estimate are the same.
 - c. The sample average has the lowest variance among all linear estimators.
 - d. All of the above. (a, b, and c are false).
 - e. None of the above (a, b, and c are not false).

8. Considere el siguiente modelo estimado:

$$Y_i = \beta_0 + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \varepsilon_i,$$

donde Y_i representa el salario del i -ésimo graduado de la Universidad (en millones de pesos). Además:

$$X_{2i} = \begin{cases} 0 & \text{si el } i\text{-ésimo graduado tiene un título de posgrado.} \\ 1 & \text{ow} \end{cases}$$

y X_{2i} es el número de años trabajados desde la graduación de la de la Universidad.

Suponiendo que la experiencia aumenta el salario:

- a. Esperamos que β_2 sea cero.
- b. Esperamos que β_2 sea positivo y mayor que β_1 .
- c. Esperamos que β_2 sea positivo.
- d. Esperamos que β_2 sea negativo
- e. Ninguna de las anteriores.

(Questions 9, 10, and 11 use the following regression results, standard errors in parentheses, $n = 200$).

$$Q_d = -500 - 100P_a + 50P_b + .3I + .2A$$

(250) (50) (30) (.1) (.08)

$$R^2 = .12$$

Where Q_d = quantity demanded of product "A", P_a = price of product "A", P_b = price of product "B" (competitor's price), I = per capita income, A = monthly advertising expenditure

9. Which of the variables is **not significant** (at the .05 level of significance)?
- a. P_a
 - b. P_b
 - c. A
 - d. I
 - e. None of the above

10. As a researcher, which aspect of the results would be of greatest concern?
- a. The negative value of the intercept.
 - b. The relatively low impact of the competitor's price
 - c. The poor fit of the regression line
 - d. The fact that the price of product "A" is not statistically significant
 - e. The fact that the good is a necessary good

11. As the manager of Product A, which of the following would be of greatest concern (based on the regression results above)?
- a. None of the factors below would be of concern.
 - b. The up-coming recession.
 - c. Pressure on you by your salespersons to lower the price so that they can boost their sales.
 - d. A price increase by the makers of product B.
 - e. Global warming

12. Considere una variable aleatoria $X \sim N(\mu, \sigma^2)$, con $\mu \neq 0$ y $\sigma^2 \neq 1$. $N(0,0, 1,0)$. Además considere la siguiente

transformación de la variable aleatoria X , $Q = (X - \mu) / \sigma$. Entonces:

- a. La varianza de Q es σ
- b. La varianza de Q es σ^2 .
- c. La varianza de Q es $1 * \sigma$
- d. La varianza de Q es $-1 * \sigma^2$
- e. Ninguna de las anteriores

13. Let X and Y be two random variables. They are such that any set of joint observations of X and Y will lie along a straight line. No matter how many observations you draw, they will always lie along this straight line. Let ρ_{XY} be the correlation coefficient between X and Y , and R^2 and $\hat{\beta}$ come from the following OLS regression using a sample of joint observations on Y and X :

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Which of the following may NOT be true?

- a. $\rho_{XY} = -1$
- b. $R^2 = 1$
- c. $\hat{\beta} = -1$
- d. $\rho_{XY} = 1$
- e. $SST = SSR + SSE$

14. Un econometrista muy joven recuerda de su curso de Econometría que uno de los siguientes modelos Probit puede ser estimado por el Método de Máxima Verosimilitud. ¿Cuál?

- a. $Y_i = \Phi(X_i^T \beta^T) + \varepsilon_i$
- b. $Y_i = \Phi(\beta^T X_i^T) \varepsilon_i$
- c. $Y_i = \Phi(\beta^T X_i^T) + \varepsilon_i$
- d. $Y_i = \Phi(\beta^T X^T) + \varepsilon_i$

e. Ninguna de las anteriores

15. When a linear regression model excludes a constant term, which of the following equalities is not correct?

- a. $SST = SSE + SSR$
- b. $R^2 = SSR/SST$

- c. $R^2 = 1 - SSE/SST$
- d. Adjusted $R^2 = 1 - (1 - R^2) (n-k)/(n-1)$

16. La significancia estadística de un parámetro en un modelo de regresión lineal se refiere a:

a. El rechazo de la hipótesis nula de que dicho parámetro es igual a cero a favor de la alternativa de que es distinto de cero.

b. El aceptar la hipótesis nula de que dicho parámetro es igual a cero en contra de la alternativa de que es distinto de cero.

c. El rechazo de la hipótesis nula de que dicho parámetro es estadísticamente igual a uno a favor de la alternativa de que es estadísticamente distinto de uno.

d. El rechazo de la hipótesis nula de que dicho parámetro es estadísticamente igual a uno a favor de la alternativa de que es estadísticamente distinto de uno.

e. Ninguna de las anteriores.

17. Which of the following would be an example of a discrete random variable:

- a. The monthly electric bill for a local business.
- b. The number of people eating at a local café between noon and 2:00 p.m.
- c. The amount of time it takes for a worker to complete a complex task.
- d. The percentage of people living below the poverty level in Colombia.
- e. None of the above.

18. La existencia de heterocedasticidad es un problema porque:

- a. En su forma habitual, MCO supone homocedasticidad y calcula las estimaciones de los parámetros del modelo con base a ese supuesto erróneo.

- b. En su forma habitual, MCO supone homocedasticidad y calcula las estimaciones de la varianza del estimador con base a ese supuesto erróneo.
- c. Sesga el estimador MCO de los parámetros.
- d. No sabemos cuál es la varianza del estimador MCO.
- e. Ninguna de las anteriores
19. The following will not cause correlation between X and u in the simple regression model:
- a. simultaneous causality.
- b. omitted variables.
- c. irrelevance of the regressor.
- d. errors in variables.
- e. None of the above
20. En el siguiente modelo $Y_t = \beta_1 + \beta_2 X_t + \mu_t$ (t=1,...N), las perturbaciones presentan heteroscedasticidad si:
- a. $cov(X_t * \mu_t) = 0$, para todo t=1,...N.
- b. $var(\mu_t) = N * \sigma^2$, para todo t=1,...N.
- c. $var(\mu_t) = c \cdot \sigma^2$, para todo t=1,...N.
- d. $\mu_t = 10 + \varepsilon_t$, con $var(\varepsilon_t) = 5$, para todo t=1,...N.
- e. Ninguna de las anteriores
21. To test whether or not the population regression function is linear rather than a polynomial of order r,
- a. check whether the regression R^2 for the polynomial regression is higher than that of the linear regression.
- b. compare the SST from both regressions.
- c. look at the pattern of the coefficients: if they change from positive to negative to positive, etc., then the polynomial regression should be used.
- d. use the test of (r-1) restrictions using the F-statistic.
- e. All of the above.
22. Si encontramos que una ecuación está sub-identificada en un sistema de ecuaciones simultáneas, entonces:
- a. No podemos estimar los parámetros de forma reducida.
- b. No podemos estimar los coeficientes de tal forma que eliminemos el problema de simultaneidad.
- c. Podemos encontrar los parámetros estructurales.
- d. No podemos estimar los parámetros por el método de MC2E.
- e. Ninguna de las anteriores.
23. A survey of earnings contains an unusually high fraction of individuals who state their weekly earnings in 100s, such as 300, 400, 500, etc. This is an example of
- a. errors-in-variables bias.
- b. variable-omission bias.
- c. simultaneous causality bias.
- d. companies that typically bargain with workers in 100s of dollars.
- e. All of the above.
24. Si en un modelo de regresión lineal múltiple el siguiente supuesto es correcto: "Las X's son estocásticas y linealmente independientes entre sí". Entonces tenemos lo que se conoce como:
- a. Heteroscedasticidad
- b. Multicolinealidad
- c. Autocorrelación
- d. Otros problemas econométricos
- e. Ninguno de los anteriores
25. $cov(u_t, u_s | X) = 0$ for $t \neq s$ means that
- a. there is no perfect multicollinearity in the errors.
- b. division of errors by regressors in different time periods is always zero.
- c. there is no correlation over time in the residuals.

- d. conditional on the regressors, the errors are uncorrelated over time.
 e. All of the above.
26. En la tabla Anova, la variación no explicada por la regresión la podemos calcular a partir de la siguiente fórmula:
 a) $\sum_{i=1}^n (y_i - \hat{y}_i)^2$
 b) $\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2$
 c) $\sum_{i=1}^n (y_i - \bar{y}_i)^2$
 d) Todas las anteriores
 e) Sólo las opciones (a) y (b) son correctas.
27. In the linear probability model, the interpretation of the slope coefficient is
 a. the change in odds associated with a unit change in X, holding other regressors constant.
 b. not all that meaningful since the dependent variable is either 0 or 1.
 c. the change in probability that $Y=1$ associated with a unit change in X, holding others regressors constant.
 d. the response in the dependent variable to a percentage change in the regressor.
 e. a. and b. are correct.
28. En un sistema de ecuaciones, cuando la ecuación uno presenta $k=2$, $g=4$ y la ecuación dos presenta $k=3$ $g=6$. Podemos afirmar entonces que:
 a. Ecuación uno y dos están sobreidentificada
 b. Ecuación uno y dos están perfectamente identificada
 c. Ecuación uno y dos están subidentificada
 d. No se puede afirmar nada
 e. Ninguna de las anteriores
29. In the probit model $P(Y_i = 1 | \mathbf{x}_i) = \Phi(\beta^T \mathbf{x}_i)$
 a. the β 's do not have a simple interpretation.
 b. the β 's tell you the effect of a unit increase in X on the probability of Y.
- c. β_0 cannot be negative since probabilities have to lie between 0 and 1.
 d. β_0 is the probability of observing Y when all X's are 0.
 e. a and c are correct.
30. Si el "p-value", asociado a la prueba de Normalidad de Jarque-Bera ($\alpha = 0.05$), toma el valor de 0.5, podemos concluir a partir de dicho estadístico que:
 a. Se rechaza la Hipótesis nula de Normalidad
 b. Se rechaza la hipótesis nula de No Normalidad
 c. No existe evidencia para rechazar la hipótesis nula de Normalidad
 d. No existe evidencia para rechazar la hipótesis nula de No Normalidad
 e. Ninguna de las anteriores.
31. The slope estimator, β_3 , has a smaller standard error, other things equal, if
 a. there is less variation in the explanatory variable, X_3 .
 b. there is more variation in the explanatory variable, X_3 .
 c. there is a large variance of the error term.
 d. the sample size is smaller.
 e. the intercept, β_0 , is smaller.
32. Si el p-valor de la prueba de Durbin-Watson para probar $H_0 : \rho = 0$ es 0.17 y el p-valor del DW para probar $H_0 : \rho \geq 0$ es 0.08, con un 10% de significancia, y según para lo que la prueba ha sido diseñada, podemos concluir de ésta que:
 a. Hay autocorrelación positiva de primer grado.
 b. Hay autocorrelación negativa de primer grado.
 c. Nos encontramos con un proceso autoregresivo de un proceso superior a 0.

- d. Nos encontramos con un proceso autoregresivo de un proceso inferior a 0.
- e. Ninguna de las anteriores.
33. When there are omitted variables in the regression, which are determinants of the dependent variable, then
- you cannot measure the effect of the omitted variable, but the estimator of your included variable(s) is (are) unaffected.
 - this has no effect on the estimator of your included variable because the other variable is not included.
 - this will always bias the OLS estimator of the included variable.
 - the OLS estimator is biased if the omitted variable is correlated with the included variable and may be possible to know the sign of the bias.
 - the OLS estimator is biased if the omitted variable is correlated with the included variable and may be *impossible* to know the sign of the bias.
34. Una deficiencia importante, entre otras, de los Modelos de Probabilidad Lineal, cuando se le compara con los Modelos Logit y Probit es:
- El permitir valores pronosticados en el intervalo entre [0, 1].
 - La Función de Distribución de probabilidad que utilizan.
 - La Probabilidad varía linealmente, en la misma proporción constante, ante un cambio unitario de la variable independiente.
 - Los coeficientes estimados no tienen una interpretación intuitiva.
 - Ninguna de las anteriores
35. If the estimates of the coefficients of interest change substantially across different specifications,
- then this can be expected from sample variation.
 - then you should change the scale of the variables to make the changes appear to be smaller.
 - then this often provides evidence that there may be a problem with the sample.
 - then choose the specification for which your coefficient of interest is most significant.
 - None of the above.
36. En un sistema de ecuaciones simultáneas, si no existe un problema de simultaneidad, entonces:
- Los MCO son eficientes y consistentes mientras que los EMC2E serán consistentes, pero no eficientes.
 - Los MCO están sesgados mientras que los EMC2E serán consistentes, pero no eficientes.
 - Los MCO son consistentes pero no eficientes, mientras que los EMC2E son eficientes y consistentes.
 - Los MCO son eficientes y consistentes mientras que los EMC2E serán inconsistentes, pero eficientes.
 - Ninguna de las anteriores.
37. You have estimated a logit model to determine the probability that an individual is earning more than twenty dollars an hour, with observations earning more than twenty dollars an hour coded as ones; You have found that $\hat{\beta}^T \mathbf{x}_i^T$ is:
- $$-22 + 2*Ed - 6*Female + 4*Exp$$
- where Ed is years of education, Female is a dummy with value one for females, and Exp is years of experience. In this case 4 is called
- the slope estimator
 - the slope estimate
 - the marginal probability
 - a and b.
 - None of the above
38. Es un supuesto del test de DW, para probar

- correlación serial:
- El modelo de regresión no incluye una constante.
 - La autocorrelación sigue un proceso AR(1) y AR(2)
 - La autocorrelación sigue un proceso AR(p), donde $p > 1$
 - La ecuación incluye una variable dependiente rezagada como regresor.
 - Ninguna de las anteriores
39. Continuing with the same information of the question e, suppose you believe that the influence of experience is a function of gender and Education. To incorporate this into your Logit model you should
- add an interaction variable defined as the product of Ed and Female
 - estimate using only the female observations and again using only the male observations
 - add a new explanatory variable coded as zero for the male observations and whatever is the value of the experience variable for the female observations
 - add a new explanatory variable coded as zero for the male observations and whatever is the value of the education times the experience variable for the female observations
 - none of the above
40. Consider two different linear estimators, $\hat{\beta}$ and $\tilde{\beta}$, of a population parameter β from a linear regression model. Suppose $E(\hat{\beta}) = \beta$, $E(\tilde{\beta}) = \beta$, and $\text{Var}(\hat{\beta}) < \text{Var}(\tilde{\beta})$. Then, all else equal
- $\hat{\beta}$ has to be the maximum likelihood estimator of β .
 - $\tilde{\beta}$ is an unambiguously better estimator than $\hat{\beta}$.
 - $\hat{\beta}$ has to be the OLS estimator.
 - a. and c. are true.
 - None of the above
41. Bajo el cumplimiento estricto de los supuestos del teorema de Gauss-Markov, no se puede decir de la matriz $(X^T X)^{-1}$ que:
- Las entradas de la diagonal toman valores estrictamente no negativos.
 - Los elementos fuera de la diagonal pueden ser negativos.
 - Es una matriz de rango completo.
 - Las entradas de la diagonal son iguales.
 - Ninguna de las anteriores.
42. The OLS residuals ($\hat{\varepsilon}_i$)
- are unbiased estimators of ε_i .
 - are unobservable since we are not able to observe the population.
 - should not be used in practice, since they indicate that your regression does not run through all your observations.
 - are MELI thanks to the Gauss-Markov theorem.
 - None of the above
43. Si una variable empleada en un modelo de regresión presenta un error de medición, entonces:
- Los estimadores MCO de los coeficientes son siempre insesgados.
 - Los estimadores MCO de los coeficientes no siempre serán sesgados.
 - Los estimadores MCO de los coeficientes siempre serán consistentes.
 - Los estimadores MCO de los coeficientes siempre serán inconsistentes.
 - Ninguna las anteriores
44. When there are omitted variables in the regression, which are not determinants of the dependent variable, then
- you cannot measure the effect of the omitted variable, but the estimator of your included variables are unaffected.
 - this has no effect on the estimates of your included variables because the other variable is not included.

- c. this will always bias the OLS estimator of the included variables.
- d. the OLS estimator is biased if the omitted variable is correlated with at least one of the included variables.
- e. We will have the famous “simultaneity bias”.
45. En un contraste de hipótesis basado en el estadístico t , si el valor absoluto de este estadístico es superior al valor crítico de la distribución:
- Podemos decir que los resultados de la regresión estimada son significativos
 - Podemos rechazar la hipótesis de que los errores son homocedásticos
 - Podemos rechazar la hipótesis nula
 - Podemos decir que la mayoría de los valores observados de la variable dependiente están cerca de la recta de regresión
 - Ninguna de las anteriores
46. A microeconomist wants to determine how corporate sales are influenced by capital and wage spending by large companies. A statistical analyst discovers that capital spending has a significant inverse relationship with corporate sales. What should this microeconomist be particularly concerned about in developing a multiple regression model using 200 observations of medium size enterprises of the same industry?
- Autocorrelation.
 - Collinearity.
 - Normality of error terms.
 - Heteroscedasticity.
 - b, c and d.
47. En el modelo de regresión múltiple, el R^2 -Ajustado:
- No puede ser negativo
 - Nunca es mayor que el coeficiente R^2
 - Es el cuadrado del coeficiente de correlación X
 - No puede disminuir cuando una nueva variable explicativa se añade al modelo
- e. Ninguna de las anteriores
48. ¿Cuál de las siguientes hipótesis no puede contrastarse mediante un estadístico F :
- $\beta_2 = 1, \beta_3 = \frac{\beta_4}{\beta_5}$
 - $\beta_2 = 0$
 - $\beta_1 + \beta_2 = 1, \beta_3 = -2\beta_4$
 - $\beta_0 = \beta_1, \beta_1 = 0$
 - Ninguna de las anteriores
49. Cuando existe un problema de variables omitidas, la condición $E(u_i | X_1, \dots, X_n) = 0$, no se satisface. Esto implica que:
- La suma de los residuos no es cero
 - El estimador de Mínimos Cuadrados Ponderados (MCP) es MELI
 - Existe un problema de multicolinealidad
 - El estimador de MCO no es consistente
 - Ninguna de las anteriores
50. En el modelo de regresión simple, la media muestral de los residuos MCO es:
- Un número positivo, puesto que en la estimación MCO se utilizan cuadrados
 - Cero
 - Inobservable, puesto que la función de regresión poblacional es desconocida
 - Su signo depende de si la variable explicativa es positiva o negativa
 - Ninguna de las anteriores

II. (25 puntos)

El economista jefe del Banco de Fomento Planetario está interesado en estimar un modelo econométrico para estudiar si la apertura de la economía ($open_t$) llevó definitivamente a menores tasas de inflación (inf_t) en la República Caribeña de los Santos. Para lo cual propone el siguiente modelo donde la tasa de inflación (inf_t) depende de la apertura de la economía ($open_t$) medida como la balanza comercial (como porcentaje del PIB) en el periodo t , y del logaritmo de la renta per cápita medido en billones de pesos en el periodo t ($lnpcinc_t$); mientras que la apertura de la economía ($open_t$) depende de la tasa de inflación (inf_t), del logaritmo de la renta per cápita ($lnpcinc_t$) y del logaritmo de la tierra de cultivo ($ln\ land_t$) medio en número de hectáreas.

- Escriba correctamente el modelo descrito en el enunciado y explique la posible existencia de un problema en los EMCO derivado de la especificación del modelo. **(4 puntos)**
- Explique claramente con qué métodos se puede estimar cada una de las expresiones **(4 puntos)**.
- A partir de la forma reducida del sistema, muestre como se puede encontrar el coeficiente que acompaña a $open_t$ en la ecuación estructural, si se emplea el método de mínimos cuadrados indirectos. **(7 puntos)**.
- Teniendo en cuenta los resultados presentados en la tabla 1, interprete los coeficientes teniendo en cuenta la significancia **(5 puntos)**.
- Después de terminado el estudio, el Presidente del Banco le envía el estudio al ministro de hacienda. Los técnicos de hacienda afirman que el estudio no es válido porque se debía emplear el método de MCO y no MC2E. ¿Cómo se puede determinar quien tiene la razón, los técnicos de hacienda o los del banco? Sea lo más claro posible en su respuesta. **(5 puntos)**

III. (25 puntos)

Un investigador está interesado en estimar los determinantes del ingreso de los minifundistas productores de caña de azúcar en la región suroccidental de una República caribeña. Para lograr su objetivo, el investigador desea emplear el siguiente modelo:

$$income_i = \beta_1 + \beta_2 \frac{1}{area_i} + \beta_3 trab_i + \beta_4 trab_i^2 + \mu_i \quad (1)$$

Donde $area_i$ es el área cultivada (en m^2) en el minifundio i , $trab_i$ es el número de trabajadores que se emplean en el minifundio i . Lastimosamente no existe una buena base de datos disponible que permita medir el ingreso de los latifundios en esa pequeña República.

No obstante, el investigador cree que puede emplear el hecho que los minifundistas reciben un subsidio del gobierno nacional por medio del programa Agro Ingreso Seguro, después de un minucioso, el minifundista recibe un subsidio si su ingreso no es superior a 100 millones de patacones al año. Afortunadamente, el listado de receptores de los subsidios está disponible en la página web y permite crear la variable $income_i$. Esta variable toma el valor de uno si el minifundista recibe un subsidio del programa Agro Ingreso Seguro o no.

- a) Discuta que tipo de problema(s) econométrico(s) puede esperar (a priori) el investigador si estima (1) por MCO. **Sea lo más concreto y preciso posible (4 puntos)**

El investigador realiza las transformaciones necesarias para obtener estimadores MELI si emplea el método de MCO a partir de las nuevas matrices $X^T X$ y $X^T y$. El asistente del investigador preparó las siguientes matrices:

$$X^T X = \begin{bmatrix} 25 & 20 & 20 & 0 \\ 20 & 30 & 20 & 0 \\ 20 & 20 & 20 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix} \quad X^T y = \begin{bmatrix} 50 \\ 5 \\ 20 \\ 10 \end{bmatrix}$$

- b) Explique claramente a que corresponden los elementos (1,1), (1,2), (2,2), (3,3) y (3,1) de la matriz $X^T X$. **Expresé su respuesta en términos de las variables en el modelo original. (5 puntos, 1 punto cada uno) (8 puntos)**
- c) Encuentre los estimadores MCO del modelo (1). **(5 Puntos)**
- d) Interprete el significado de cada uno de los coeficientes estimados. **(8 Puntos, 2 puntos cada uno)**

Tabla 1. Resultados de EasyReg.

Two-stage least squares:		
Dependent variable:		
Y = inf		
X variables, including instrumental variables:		
X(1) = open		
X(2) = lnpcinc		
X(3) = inland		
X(4) = 1		
WARNING: The effective degrees of freedom is only 12.		
Therefore, the estimation results may be unreliable!		
Endogenous X variables:		
Y*(1)=open		
Exogenous X variables:		
X*(1)=lnpcinc		
X*(2)=1		
2SLS estimation results for Y = inf		
Parameters	Estimate	t-value (S.E.) [p-value]
Open	-0.3374	-26.693 (0.14166) [0.00000]
lnpcinc	0.3758	0.9650

		(0.03775)
		[0.70000]
1	26.8993	0.37650
		(0.03775)
		[0.82500]

[The p-values are two-sided and based on the normal approximation]

Standard error of the residuals = 1.2961
 Residual sum of squares (RSS) = 60,64
 Total sum of squares (TSS) = 606,4

R-squared = 0.81

Effective sample size (n) = 150
 DW= 1.06

Econometría 06216
Examen Final
Respuestas Sugeridas
Cali, Miércoles 17 de noviembre de 2010

Profesores: Julio César Alonso --- Carlos Giovanni González

Estudiante: _____

Código: _____

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **12** páginas; además, deben tener 2 páginas de fórmulas.
3. El examen consta de 3 preguntas que suman un total de 100 puntos. El valor de cada una de las preguntas esta 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.
5. El examen está diseñado para dos horas, pero ustedes tienen 4 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 respuestas escritas en 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!

I. Selección Múltiple (50 puntos en total, 1 punto por cada subparte)

Seleccione la opción **más indicada** en la hoja de respuestas que encontrará al final de este examen. Sólo se considerarán respuestas que sean consignadas en la hoja de respuestas. (No es necesario justificar su respuesta)

1. Suppose the following model $y_i = \beta_1 + \beta_2 D_i x_{1i} + \beta_3 W_i x_{1i} + \varepsilon_i$ where D_i and W_i are dummy variables, and $D_i = \begin{cases} 1 & \text{female} \\ 0 & \text{o.w} \end{cases}$ and $W_i = \begin{cases} 1 & \text{male} \\ 0 & \text{o.w} \end{cases}$, therefore the β 's estimates will most likely be:
 - a. Inconsistent.
 - b. Perfectly multicollinear.
 - c. Unbiased.
 - d. Biased.
 - e. Efficient.
2. Un econométra está convencido de la necesidad de una de los siguientes supuestos para que el estimador MCO de β sea insesgado. ¿Cuál es ese supuesto?:
 - a) Que la varianza del término de error sea constante a lo largo de la muestra
 - b) Que la distribución de los errores sea normal
 - c) Que los parámetros del modelo no sean constantes a lo largo de la muestra
 - d) Que la covarianza entre cualquier par de errores sea igual a cero
 - e) Ninguno de los anteriores
3. One tailed tests are *sometimes* used to test hypotheses about regression coefficients. In which of the following circumstances?
 - a. when the estimated coefficient is close to zero

- b. when you wish to use a larger significance level than 5%
 - c. when the sample size is large enough to use the normal approximation to the t distribution
 - d. when the estimated coefficient has the opposite sign to that predicted by theory
 - e. when you are testing a hypothesis other than that the parameter equals zero
4. Considere el siguiente modelo de regresión $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$; donde, μ cumple los supuestos habituales. Y con respecto a las variables originales Y , X_1 , X_2 , el modelo es:
 - a. Lineal en parámetros y en variables
 - b. No lineal en parámetros y lineal en variables
 - c. Lineal en parámetros y no lineal en variables
 - d. No lineal en parámetros y no lineal en variables
 - e. Ninguna de las anteriores
 5. The fact that least squares estimator of the slope coefficients is unbiased means that:
 - a. the estimated slope coefficient will always be equal to the true parameter value
 - b. the estimated slope coefficient will get closer to the true parameter value as the size of the sample increases
 - c. the estimated slope coefficient will be equal to the true parameter if the sample is large
 - d. the mean of the sampling distribution of the slope parameter is zero
 - e. if repeated samples of the same size are taken, on average their value will be equal to the true parameter
 6. Un econométra experimentado afirma que en uno de los siguientes modelos los parámetros β_1 y β_2 no podrían ser estimados por MCO (donde μ es el término de error). ¿Cuál?
 - a. $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i$
 - b. $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i + \mu_i$
 - c. $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i + \mu_i^2$
 - d. $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \mu_i + \mu_i^3$

- a) $Y_i = \beta_1 + \beta_2 (\ln X_i)^2 + \mu_i$
- b) $Y_i = e^{(\beta_1 + \beta_2 X_i + \mu_i)}$
- c) $Y_i = \sqrt{\beta_1 + \beta_2 X_i} + \mu_i$
- d) $Y_i = e^{\beta_1} X_i^{\beta_2} e^{\mu_i}$
- e) Ninguna de las anteriores

7. Consider the classical econometrics model with noexplanatory variables ($n > 1$), $Y_i = \beta_1 + \beta_2 X_i + \mu_i$. Which of the following statements is false?
- a. The sample average is a linear estimator.
 - b. The sample average and the OLS estimate are the same.
 - c. The sample average has the lowest variance among all linear estimators.
 - d. All of the above. (a, b, and c are false).
 - e. None of the above (a, b, and c are not false).

8. Considere el siguiente modelo estimado:

$$Y_i = \beta_0 + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \varepsilon_i$$

donde Y_i representa el salario del i -ésimo graduado de la Universidad (en millones de pesos). Además:

$$X_{1i} = \begin{cases} 0 & \text{si el } i\text{-ésimo graduado tiene un título de posgrado.} \\ 1 & \text{ow} \end{cases}$$

y X_{2i} es el número de años trabajados desde la graduación de la de la Universidad.

Suponiendo que la experiencia aumenta el salario:

- a. Esperamos que β_2 sea cero.
- b. Esperamos que β_2 sea positivo y mayor que β_1 .
- c. Esperamos que β_2 sea positivo.
- d. Esperamos que β_2 sea negativo
- e. Ninguna de las anteriores.

(Questions 9, 10, and 11 use the following regression results, standard errors in parentheses, $n = 200$).

$$Q_d = -500 - 100P_a + 50P_b + .3I + .2A$$

(250) (50) (30) (.1) (.08)

$$R^2 = .12$$

Where Q_d = quantity demanded of product "A", P_a = price of product "A", P_b = price of product "B" (competitor's price), I = per capita income, A = monthly advertising expenditure

9. Which of the variables is **not significant** (at the .05 level of significance)?
- a. P_a
 - b. P_b
 - c. A
 - d. I
 - e. None of the above

10. As a researcher, which aspect of the results would be of greatest concern?

- a. The negative value of the intercept.
- b. The relatively low impact of the competitor's price
- c. **The poor fit of the regression line**
- d. The fact that the price of product "A" is not statistically significant
- e. The fact that the good is a necessary good

11. As the manager of Product A, which of the following would be of greatest concern (based on the regression results above)?

- a. None of the factors below would be of concern.
- b. **The up-coming recession.**
- c. Pressure on you by your salespersons to lower the price so that they can boost their sales.
- d. A price increase by the makers of product B.
- e. Global warming

12. Considere una variable aleatoria $X \sim N(\mu, \sigma^2)$, con $\mu \neq 0$ y $\sigma^2 \neq 1$. $N(0, 1.0)$. Además considere la siguiente

transformación de la variable aleatoria X , $Q = (X - \mu) / \sigma$. Entonces:

- a. La varianza de Q es σ
- b. La varianza de Q es σ^2 .
- c. La varianza de Q es $1 * \sigma$
- d. La varianza de Q es $1 * \sigma^2$
- e. **Ninguna de las anteriores**

13. Let X and Y be two random variables. They are such that any set of joint observations of X and Y will lie along a straight line. No matter how many observations you draw, they will always lie along this straight line. Let ρ_{XY} be the correlation coefficient between X and Y , and R^2 and $\hat{\beta}$ come from the following OLS regression using a sample of joint observations on Y and X :

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Which of the following may NOT be true?

- a. $\rho_{XY} = -1$
- b. $R^2 = 1$
- c. **$\hat{\beta} = -1$**
- d. $\rho_{XY} = 1$
- e. $SST = SSR + SSE$

14. Un econometrista muy joven recuerda de su curso de Econometría que uno de los siguientes modelos Probit puede ser estimado por el Método de Máxima Verosimilitud. ¿Cuál?

- a. $Y_i = \Phi(X_i^T \beta^T) + \varepsilon_i$
- b. $Y_i = \Phi(\beta^T X_i^T) \varepsilon_i$
- c. **$Y_i = \Phi(\beta^T X_i^T) + \varepsilon_i$**
- d. $Y_i = \Phi(\beta^T X^T) + \varepsilon_i$
- e. Ninguna de las anteriores

15. When a linear regression model excludes a constant term, which of the following equalities is not correct?

- a. **$SST = SSE + SSR$**
- b. $R^2 = SSR/SST$

- c. $R^2 = 1 - SSE/SST$
- d. **Adjusted $R^2 = 1 - (1 - R^2) (n-k)/(n-1)$**

16. La significancia estadística de un parámetro en un modelo de regresión lineal se refiere a:

a. El rechazo de la hipótesis nula de que dicho parámetro es igual a cero a favor de la alternativa de que es distinto de cero.

b. El aceptar la hipótesis nula de que dicho parámetro es igual a cero en contra de la alternativa de que es distinto de cero.

c. El rechazo de la hipótesis nula de que dicho parámetro es estadísticamente igual a uno a favor de la alternativa de que es estadísticamente distinto de uno.

d. El rechazo de la hipótesis nula de que dicho parámetro es estadísticamente igual a uno a favor de la alternativa de que es estadísticamente distinto de uno.

e. Ninguna de las anteriores.

17. Which of the following would be an example of a discrete random variable:

- a. The monthly electric bill for a local business.
- b. **The number of people eating at a local café between noon and 2:00 p.m.**
- c. The amount of time it takes for a worker to complete a complex task.
- d. The percentage of people living below the poverty level in Colombia.
- e. None of the above.

18. La existencia de heterocedasticidad es un problema porque:

- a. En su forma habitual, MCO supone homocedasticidad y calcula las estimaciones de los parámetros del modelo con base a ese supuesto erróneo.

- b. En su forma habitual, MCO supone homocedasticidad y calcula las estimaciones de la varianza del estimador con base a ese supuesto erróneo.
 - c. Sesga el estimador MCO de los parámetros.
 - d. No sabemos cuál es la varianza del estimador MCO.
 - e. Ninguna de las anteriores
19. The following will not cause correlation between X and u in the simple regression model:
- a. simultaneous causality.
 - b. omitted variables.
 - c. irrelevance of the regressor.
 - d. errors in variables.
 - e. None of the above
20. En el siguiente modelo $Y_t = \beta_1 + \beta_2 X_t + \mu_t$ (t=1,...N), las perturbaciones presentan heteroscedasticidad si:
- a. $cov(X_t * \mu_t) = 0$, para todo t=1,...N.
 - b. $var(\mu_t) = N * \sigma^2$, para todo t=1,...N.
 - c. $var(\mu_t) = c * \sigma^2$, para todo t=1,...N.
 - d. $\mu_t = 10 + \varepsilon_t$, con $var(\varepsilon_t) = 5$, para todo t=1,...N.
 - e. Ninguna de las anteriores
21. To test whether or not the population regression function is linear rather than a polynomial of order r,
- a. check whether the regression R² for the polynomial regression is higher than that of the linear regression.
 - b. compare the SST from both regressions.
 - c. look at the pattern of the coefficients: if they change from positive to negative to positive, etc., then the polynomial regression should be used.
 - d. use the test of (r-1) restrictions using the F-statistic.
- e. All of the above.
22. Si encontramos que una ecuación está sub-identificada en un sistema de ecuaciones simultáneas, entonces:
- a. No podemos estimar los parámetros de forma reducida.
 - b. No podemos estimar los coeficientes de tal forma que eliminemos el problema de simultaneidad.
 - c. Podemos encontrar los parámetros estructurales.
 - d. No podemos estimar los parámetros por el método de MC2E.
 - e. Ninguna de las anteriores.
23. A survey of earnings contains an unusually high fraction of individuals who state their weekly earnings in 100s, such as 300, 400, 500, etc. This is an example of
- a. errors-in-variables bias.
 - b. variable-omission bias.
 - c. simultaneous causality bias.
 - d. companies that typically bargain with workers in 100s of dollars.
 - e. All of the above.
24. Si en un modelo de regresión lineal múltiple el siguiente supuesto es correcto: "Las X's son estocásticas y linealmente independientes entre sí". Entonces tenemos lo que se conoce como:
- a. Heteroscedasticidad
 - b. Multicolinealidad
 - c. Autocorrelación
 - d. Otros problemas econométricos
 - e. Ninguno de los anteriores
25. $cov(u_t, u_s | X) = 0$ for t≠s means that
- a. there is no perfect multicollinearity in the errors.
 - b. division of errors by regressors in different time periods is always zero.
 - c. there is no correlation over time in the residuals.

- d. conditional on the regressors, the errors are uncorrelated over time.
 - e. All of the above.
26. En la tabla Anova, la variación no explicada por la regresión la podemos calcular a partir de la siguiente fórmula:
- a) $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$
 - b) $\sum_{i=1}^n (\hat{Y}_i - \bar{Y}_i)^2$
 - c) $\sum_{i=1}^n (Y_i - \bar{Y}_i)^2$
 - d) Todas las anteriores
 - e) Sólo las opciones (a) y (b) son correctas.
27. In the linear probability model, the interpretation of the slope coefficient is
- a. the change in odds associated with a unit change in X, holding other regressors constant.
 - b. not all that meaningful since the dependent variable is either 0 or 1.
 - c. the change in probability that Y=1 associated with a unit change in X, holding others regressors constant.
 - d. the response in the dependent variable to a percentage change in the regressor.
 - e. a. and b. are correct.
28. En un sistema de ecuaciones, cuando la ecuación uno presenta k=2, g=4 y la ecuación dos presenta k=3 g=6. Podemos afirmar entonces que:
- a. Ecuación uno y dos están sobreidentificada
 - b. Ecuación uno y dos están perfectamente identificada
 - c. Ecuación uno y dos están subidentificada
 - d. No se puede afirmar nada
 - e. Ninguna de las anteriores
29. In the probit model $P(Y_i = 1 | x_i) = \Phi(\beta' x_i)$
- a. the β 's do not have a simple interpretation.
 - b. the β 's tell you the effect of a unit increase in X on the probability of Y.
- c. β_0 cannot be negative since probabilities have to lie between 0 and 1.
 - d. β_0 is the probability of observing Y when all X's are 0.
 - e. a and c are correct.
30. Si el "p-value", asociado a la prueba de Normalidad de Jarque-Bera ($\alpha = 0.05$), toma el valor de 0.5, podemos concluir a partir de dicho estadístico que:
- a. Se rechaza la Hipótesis nula de Normalidad
 - b. Se rechaza la hipótesis nula de No Normalidad
 - c. No existe evidencia para rechazar la hipótesis nula de Normalidad
 - d. No existe evidencia para rechazar la hipótesis nula de No Normalidad
 - e. Ninguna de las anteriores.
31. The slope estimator, β_1 , has a smaller standard error, other things equal, if
- a. there is less variation in the explanatory variable, X_1 .
 - b. There is more variation in the explanatory variable, X_1 .
 - c. there is a large variance of the error term.
 - d. the sample size is smaller.
 - e. the intercept, β_0 , is smaller.
32. Si el p-valor de la prueba de Durbin-Watson para probar $H_0 : \rho = 0$ es 0.17 y el p-valor del DW para probar $H_0 : \rho \geq 0$ es 0.08, con un 10% de significancia, y según para lo que la prueba ha sido diseñada, podemos concluir de ésta que:
- a. Hay autocorrelación positiva de primer grado.
 - b. Hay autocorrelación negativa de primer grado.
 - c. Nos encontramos con un proceso autoregresivo de un proceso superior a 0.

- d. Nos encontramos con un proceso autoregresivo de un proceso inferior a 0.
 e. Ninguna de las anteriores.
33. When there are omitted variables in the regression, which are determinants of the dependent variable, then
 a. you cannot measure the effect of the omitted variable, but the estimator of your included variable(s) is (are) unaffected.
 b. this has no effect on the estimator of your included variable because the other variable is not included.
 c. this will always bias the OLS estimator of the included variable.
 d. the OLS estimator is biased if the omitted variable is correlated with the included variable and may be possible to know the sign of the bias.
 e. the OLS estimator is biased if the omitted variable is correlated with the included variable and may be impossible to know the sign of the bias.
34. Una deficiencia importante, entre otras, de los Modelos de Probabilidad Lineal, cuando se le compara con los Modelos Logit y Probit es:
 a. El permitir valores pronosticados en el intervalo entre [0, 1].
 b. La Función de Distribución de probabilidad que utilizan.
 c. La Probabilidad varía linealmente, en la misma proporción constante, ante un cambio unitario de la variable independiente.
 d. Los coeficientes estimados no tienen una interpretación intuitiva.
 e. Ninguna de las anteriores
35. If the estimates of the coefficients of interest change substantially across different specifications,
 a. then this can be expected from sample variation.
 b. then you should change the scale of the variables to make the changes appear to be smaller.
 c. then this often provides evidence that there may be a problem with the sample.
 d. then choose the specification for which your coefficient of interest is most significant.
 e. None of the above.
36. En un sistema de ecuaciones simultáneas, si no existe un problema de simultaneidad, entonces:
 a. Los MCO son eficientes y consistentes mientras que los EMC2E serán consistentes, pero no eficientes.
 b. Los MCO están sesgados mientras que los EMC2E serán consistentes, pero no eficientes.
 c. Los MCO son consistentes pero no eficientes, mientras que los EMC2E son eficientes y consistentes.
 d. Los MCO son eficientes y consistentes mientras que los EMC2E serán inconsistentes, pero eficientes.
 e. Ninguna de las anteriores.
37. You have estimated a logit model to determine the probability that an individual is earning more than twenty dollars an hour, with observations earning more than twenty dollars an hour coded as ones; You have found that $\hat{\beta}'x_i'$ is:
 $-22 + 2*Ed - 6*Female + 4*Exp$
 where Ed is years of education, Female is a dummy with value one for females, and Exp is years of experience. In this case 4 is called
 a. the slope estimator
 b. the slope estimate
 c. the marginal probability
- Esto puede ser un indicio de multicolinealidad no perfecta, lo cual no es un problema de la muestra.

- d. a and b.
 e. None of the above
38. Es un supuesto del test de DW, para probar correlación serial:
 a. El modelo de regresión no incluye una constante.
 b. La autocorrelación sigue un proceso AR(1) y AR(2)
 c. La autocorrelación sigue un proceso AR(p), donde $p > 1$
 d. La ecuación incluye una variable dependiente rezagada como regresor.
 e. Ninguna de las anteriores
39. Continuing with the same information of the question 37, suppose you believe that the influence of experience is a function of gender and Education. To incorporate this into your Logit model you should
 a. add an interaction variable defined as the product of Ed and Female
 b. estimate using only the female observations and again using only the male observations
 c. add a new explanatory variable coded as zero for the male observations and whatever is the value of the experience variable for the female observations
 d. add a new explanatory variable coded as zero for the male observations and whatever is the value of the education times the experience variable for the female observations
 e. none of the above
40. Consider two different linear estimators, $\hat{\beta}$ and $\tilde{\beta}$, of a population parameter β from a linear regression model. Suppose $E(\hat{\beta}) = \beta$, $E(\tilde{\beta}) = \beta$, and $\text{Var}(\hat{\beta}) < \text{Var}(\tilde{\beta})$. Then, all else equal
 a. $\hat{\beta}$ has to be the maximum likelihood estimator of β .
 b. $\tilde{\beta}$ is an unambiguously better estimator than $\hat{\beta}$.
- c. $\hat{\beta}$ has to be the OLS estimator.
 d. a. and c. are true.
 e. None of the above
41. Bajo el cumplimiento estricto de los supuestos del teorema de Gauss-Markov, no se puede decir de la matriz $(X'X)^{-1}$ que:
 a. Las entradas de la diagonal toman valores estrictamente no negativos.
 b. Los elementos fuera de la diagonal pueden ser negativos.
 c. Es una matriz de rango completo.
 d. Las entradas de la diagonal son iguales.
 e. Ninguna de las anteriores.
42. The OLS residuals ($\hat{\epsilon}_i$)
 a. are unbiased estimators of ϵ_i .
 b. are unobservable since we are not able to observe the population.
 c. should not be used in practice, since they indicate that your regression does not run through all your observations.
 d. are MELI thanks to the Gauss-Markov theorem.
 e. None of the above
43. Si una variable empleada en un modelo de regresión presenta un error de medición, entonces:
 a. Los estimadores MCO de los coeficientes son siempre insesgados.
 b. Los estimadores MCO de los coeficientes no siempre serán sesgados.
 c. Los estimadores MCO de los coeficientes siempre serán consistentes.
 d. Los estimadores MCO de los coeficientes siempre serán inconsistentes.
 e. Ninguna las anteriores
44. When there are omitted variables in the regression, which are not determinants of the dependent variable, then
 a. you cannot measure the effect of the omitted variable, but the estimator of your included variables are unaffected.

- b. this has no effect on the estimates of your included variables because the other variable is not included.
 - c. this will always bias the OLS estimator of the included variables.
 - d. the OLS estimator is biased if the omitted variable is correlated with at least one of the included variables.
 - e. We will have the famous "simultaneity bias".
45. En un contraste de hipótesis basado en el estadístico t , si el valor absoluto de este estadístico es superior al valor crítico de la distribución:
- a) Podemos decir que los resultados de la regresión estimada son significativos
 - b) Podemos rechazar la hipótesis de que los errores son homocedásticos
 - c) Podemos rechazar la hipótesis nula
 - d) Podemos decir que la mayoría de los valores observados de la variable dependiente están cerca de la recta de regresión
 - e) Ninguna de las anteriores
46. A microeconomist wants to determine how corporate sales are influenced by capital and wage spending by large companies. A statistical analyst discovers that capital spending has a significant inverse relationship with corporate sales. What should this microeconomist be particularly concerned about in developing a multiple regression model using 200 observations of medium size enterprises of the same industry?
- a. Autocorrelation.
 - b. Collinearity.
 - c. Normality of error terms.
 - d. Heteroscedasticity.
 - e. b, c and d.
47. En el modelo de regresión múltiple, el R^2 -Ajustado:
- a. No puede ser negativo
 - b. Nunca es mayor que el coeficiente R^2
- c. Es el cuadrado del coeficiente de correlación X
 - d. No puede disminuir cuando una nueva variable explicativa se añade al modelo
 - e. Ninguna de las anteriores
48. ¿Cuál de las siguientes hipótesis no puede contrastarse mediante un estadístico F:
- a. $\beta_2 = 1, \beta_3 = \frac{\beta_2}{\beta_5}$
 - b. $\beta_2 = 0$
 - c. $\beta_1 + \beta_2 = 1, \beta_3 = -2\beta_4$
 - d. $\beta_0 = \beta_1, \beta_1 = 0$
 - e. Ninguna de las anteriores
49. Cuando existe un problema de variables omitidas, la condición $E(u_i | X_1, \dots, X_n) = 0$, no se satisface. Esto implica que:
- a. La suma de los residuos no es cero
 - b. El estimador de Mínimos Cuadrados Ponderados (MCP) es MELI
 - c. Existe un problema de multicolinealidad
 - d. El estimador de MCO no es consistente
 - e. Ninguna de las anteriores
50. En el modelo de regresión simple, la media muestral de los residuos MCO es:
- a. Un número positivo, puesto que en la estimación MCO se utilizan cuadrados
 - b. Cero
 - c. Inobservable, puesto que la función de regresión poblacional es desconocida
 - d. Su signo depende de si la variable explicativa es positiva o negativa
 - e. Ninguna de las anteriores

II. (25 puntos)

El economista jefe del Banco de Fomento Planetario está interesado en estimar un modelo econométrico para estudiar si la apertura de la economía ($open_t$) llevó definitivamente a menores tasas de inflación (inf_t) en la República Caribeña de los Santos. Para lo cual propone el siguiente modelo donde la tasa de inflación (inf_t) depende de la apertura de la economía ($open_t$) medida como la balanza comercial (como porcentaje del PIB) en el periodo t , y del logaritmo de la renta per cápita medido en billones de pesos en el periodo t ($lnpcinc_t$); mientras que la apertura de la economía ($open_t$) depende de la tasa de inflación (inf_t), del logaritmo de la renta per cápita ($lnpcinc_t$) y del logaritmo de la tierra de cultivo ($ln land_t$) medio en número de hectáreas.

- a. Escriba correctamente el modelo descrito en el enunciado y explique la posible existencia de un problema en los EMCO derivado de la especificación del modelo. (4 puntos)

El sistema de ecuaciones descrito en el enunciado es el siguiente (2 puntos):

$$inf_t = \beta_0 + \beta_1 open_t + \beta_2 lnpcinc_t + \mu_t$$

$$open_t = \alpha_0 + \alpha_1 inf_t + \alpha_2 lnpcinc_t + \alpha_3 lnland_t + v_t$$

El posible problema es de simultaneidad, el término de error y una de las variables explicativas de una de las ecuaciones están correlacionados, y los EMCO serán sesgados e inconsistentes (2 puntos). No se dará crédito, si se mencionan los problemas de autocorrelación y heteroscedasticidad.

- b. Explique claramente con qué métodos se puede estimar cada una de las expresiones (4 puntos).

En este caso las variables endógenas son: (inf_t) y ($open_t$). Las variables exógenas son: ($ln land_t$) y ($lnpcinc_t$). En ese orden de ideas tenemos que (4 puntos, 2 puntos por cada ecuación.):

Ecuación	k_i	g_i	$k_i ? g_i - 1$	¿Sub, perfectamente o sobre identificada?	Método de estimación
Error! Reference source not found.	1	2	1=1	Perfectamente identificada (probablemente)	MCI o MC2E
Error! Reference source not found.	0	2	0<1	Subidentificada	no se puede estimar

- c. A partir de la forma reducida del sistema, muestre como se puede encontrar el coeficiente que acompaña a $open_t$ en la ecuación estructural, si se emplea el método de mínimos cuadrados indirectos. (7 puntos).

Es un sistema de ecuaciones simultáneas donde las ecuaciones estructurales son:

$$inf_t = \beta_0 + \beta_1 open_t + \beta_2 lnpcinc_t + \mu_t$$

$$open_t = \alpha_0 + \alpha_1 inf_t + \alpha_2 lnpinc_t + \alpha_3 inland_t + v_t$$

Y despejando, se obtienen las ecuaciones de forma reducida, con sus respectivos multiplicadores de impacto de corto plazo:

$$open_t = \frac{1}{1 - \alpha_1 \beta_1} \{ \alpha_0 + \alpha_1 \beta_0 + [\alpha_1 \beta_2 + \alpha_2] lnpinc_t + \alpha_3 inland_t + \alpha_1 \mu_t + v_t \}$$

$$inf_t = \frac{1}{1 - \alpha_1 \beta_1} \{ \beta_0 + \beta_1 \alpha_0 + [\beta_2 + \beta_1 \alpha_2] lnpinc_t + \beta_1 \alpha_3 inland_t + \mu_t + \beta_1 v_t \}$$

Reparametrizando se obtiene que

$$open_t = \pi_{11} + \pi_{12} lnpinc_t + \pi_{13} inland_t + \varepsilon_1$$

$$inf_t = \pi_{21} + \pi_{22} lnpinc_t + \pi_{23} inland_t + \varepsilon_2$$

Note que el coeficiente asociado a la variable $open_t$ en la ecuación estructural puede hallarse por el método de mínimos cuadrados indirectos dividiendo π_{23} entre π_{13} , dado que ambos pueden estimarse sin ninguna complicación (ambas ecuaciones reducidas están perfectamente identificadas).

Es decir:

$$\hat{\beta}_1^{MCI} = \frac{\hat{\pi}_{23}}{\hat{\pi}_{13}}$$

- d. Teniendo en cuenta los resultados presentados en la tabla 1, interprete los coeficientes teniendo en cuenta la significancia **(5 puntos)**.

$\hat{\beta}_0 = 26.8993$ no tiene interpretación económica. No es estadísticamente significativo. (1 punto)

$\hat{\beta}_1 = -0.3374$ Es significativo al 99%. Ante un aumento de un punto porcentual en la balanza comercial como porcentaje del PIB, la función de inflación disminuirá en 0.3374 puntos porcentuales. (2 puntos)

$\hat{\beta}_2 = 0.3758$ No es estadísticamente significativo. La renta per cápita no tiene efecto sobre la función de inflación (2 puntos)

- e. Después de terminado el estudio, el Presidente del Banco le envía el estudio al ministro de hacienda. Los técnicos de hacienda afirman que el estudio no es válido porque se debía emplear el método de MCO y no MC2E. ¿Cómo se puede determinar quien tiene la razón, los técnicos de hacienda o los del banco? Sea lo más claro posible en su respuesta. **(5 puntos)**

Su respuesta debería contar con los siguientes elementos:

- Determinar que todo depende de si ($open_t$) tiene un problema de simultaneidad o no (1 punto)
- Se necesita emplear la prueba de Hausman y explicar (2 puntos)
- Determinar cómo se toma la decisión en favor de hacienda o el banco de acuerdo a la decisión de la prueba. (2 puntos).

III. (25 puntos)

Un investigador está interesado en estimar los determinantes del ingreso de los minifundistas productores de caña de azúcar en la región suroccidental de una República caribeña. Para lograr su objetivo, el investigador desea emplear el siguiente modelo:

$$income_i = \beta_1 + \beta_2 \frac{1}{area_i} + \beta_3 trab_i + \beta_4 trab_i^2 + \mu_i \quad (1)$$

Donde $area_i$ es el área cultivada (en m^2) en el minifundio i , $trab_i$ es el número de trabajadores que se emplean en el minifundio i . Lastimosamente no existe una buena base de datos disponible que permita medir el ingreso de los latifundios en esa pequeña República.

No obstante, el investigador cree que puede emplear el hecho que los minifundistas reciben un subsidio del gobierno nacional por medio del programa Agro Ingreso Seguro, después de un minucioso, el minifundista recibe un subsidio si su ingreso no es superior a 100 millones de patacones al año. Afortunadamente, el listado de receptores de los subsidios está disponible en la página web y permite crear la variable $income_i$. Esta variable toma el valor de uno si el minifundista recibe un subsidio del programa Agro Ingreso Seguro y **cero si no**.

- a) Discuta que tipo de problema(s) econométrico(s) puede esperar (a priori) el investigador si estima (1) por MCO. **Sea lo más concreto y preciso posible (4 puntos)**

Noten que estimar (1) implica tener una variable dummy como variable dependiente. Así, si empleamos el método de mínimos cuadrados ordinarios para estimar este modelo tendremos un problema de heteroscedasticidad. Es decir, dado que esto corresponde al modelo de probabilidad lineal **(1 punto)** se espera que $\sigma_i^2 = income_i (1 - income_i)$. **(3 puntos)**

Noten que a priori no hay porque esperar un problema de multicolinealidad perfecta (menos 2 puntos por decir que existía este problema)

El investigador realiza las transformaciones necesarias para obtener estimadores MELI si emplea el método de MCO a partir de las nuevas matrices $X^T X$ y $X^T y$. El asistente del investigador preparó las siguientes matrices:

$$X^T X = \begin{bmatrix} 25 & 20 & 20 & 0 \\ 20 & 30 & 20 & 0 \\ 20 & 20 & 20 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix} \quad X^T y = \begin{bmatrix} 50 \\ 5 \\ 20 \\ 10 \end{bmatrix}$$

b) Explique claramente a que corresponden los elementos (1,1), (1,2), (2,2), (3,3) y (3,1) de la matriz $X^T X$. **Exprese su respuesta en términos de las variables en el modelo original. (8 puntos)**

Noten que para poder obtener estimadores MCO se tiene que emplear el método de MCP. De tal manera que el modelo transformado corresponde a:

$$income_i = \beta_1 + \beta_2 \frac{1}{area_i} + \beta_3 trab_i + \beta_4 trab_i^2 + \mu_i$$

$$income_i = \beta^T x_i + \mu_i$$

Definamos $w_i^2 = \hat{\beta}^{*T} x_i [1 - \hat{\beta}^{*T} x_i]$, donde $\hat{\beta}^*$ corresponde a los estimadores MCO del modelo (1).

Entonces el modelo transformado será:

$$\frac{income_i}{w_i} = \beta_1 \frac{1}{w_i} + \beta_2 \frac{1}{area_i} \frac{1}{w_i} + \beta_3 \frac{trab_i}{w_i} + \beta_4 \frac{trab_i^2}{w_i} + \frac{\mu_i}{w_i}$$

Sino se reconoce que implica que se emplea el valor estimado menos 2 puntos

Por tanto se tiene que:

Elemento (1,1)

$$\sum \frac{1}{w_i^2} = \sum \frac{1}{\left(\hat{\beta}^*_1 + \hat{\beta}^*_2 \frac{1}{area_i} + \hat{\beta}^*_3 trab_i + \hat{\beta}^*_4 trab_i^2 \right) \left(1 - \left(\hat{\beta}^* x_i \right) \right)}$$

Elemento (1,2)

$$\sum \frac{1}{area_i} \frac{1}{w_i^2} = \sum \frac{1}{area_i} \frac{1}{\left(\hat{\beta}^*_1 + \hat{\beta}^*_2 \frac{1}{area_i} + \hat{\beta}^*_3 trab_i + \hat{\beta}^*_4 trab_i^2 \right) \left(1 - \left(\hat{\beta}^* x_i \right) \right)}$$

Elemento (2,2)

$$\sum \frac{1}{area_i^2} \frac{1}{w_i^2} = \sum \left(\frac{1}{area_i^2} \frac{1}{\left(\hat{\beta}^*_1 + \hat{\beta}^*_2 \frac{1}{area_i} + \hat{\beta}^*_3 trab_i + \hat{\beta}^*_4 trab_i^2 \right) \left(1 - \left(\hat{\beta}^* x_i \right) \right)} \right)$$

Elemento (3,1)

$$\sum \frac{1}{w_i^2} = \sum \frac{trab_i}{\left(\hat{\beta}^*_1 + \hat{\beta}^*_2 \frac{1}{area_i} + \hat{\beta}^*_3 trab_i + \hat{\beta}^*_4 trab_i^2 \right) \left(1 - \left(\hat{\beta}^* x_i \right) \right)}$$

Elemento (3,3)

$$\sum \frac{1}{w_i^2} = \sum \left(\frac{trab_i^2}{\left(\hat{\beta}^*_1 + \hat{\beta}^*_2 \frac{1}{area_i} + \hat{\beta}^*_3 trab_i + \hat{\beta}^*_4 trab_i^2 \right) \left(1 - \left(\hat{\beta}^* x_i \right) \right)} \right)$$

c) Encuentre los estimadores MCO del modelo (1). **(5 Puntos)**

En este caso tenemos que:

$$\hat{\beta}_{hat} = (X^T X)^{-1} \cdot X^T y$$

$$\hat{\beta} = \begin{bmatrix} 1/5 & 0 & -1/5 & 0 \\ 0 & 1/10 & -1/10 & 0 \\ -1/5 & -1/10 & 7/20 & 0 \\ 0 & 0 & 0 & 1/10 \end{bmatrix} \begin{bmatrix} 50 \\ 5 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 6 \\ -3/2 \\ -7/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6.0 \\ -1.5 \\ -3.5 \\ 1.0 \end{bmatrix}$$

d) Interprete el significado de cada uno de los coeficientes estimados. **(8 Puntos, 2 puntos cada uno)**

$\hat{\beta}_1 = 6.0$. Una interpretación al pie de la letra sería: Un minifundista recibirá con seguridad un subsidio de AIS cuando se tiene un área cultivada infinita y no se tiene trabajadores. Noten que como se trata de minifundios, no tiene sentido hablar de un área que tiende a infinito. Por tanto si bien tiene una interpretación, el resultado obtenido no tiene mucho sentido.

$\hat{\beta}_2 = -1.5$ Cuando el área de un minifundio aumenta en un uno por ciento, la probabilidad de recibir un subsidio de AIS aumentará en $\frac{1.5}{area_i}$ puntos porcentuales.

NOTEN que dado que esperamos que los minifundios tengan más de un metro cuadrado, entones este aumento no necesariamente será mayor de 100 puntos porcentuales. Por tanto la interpretación anterior tiene sentido.

Dado que:

$$\frac{\partial income_i}{\partial trab_i} = \beta_3 + 2\beta_4 trab_i$$

$\hat{\beta}_3 = -3.5$ es la parte constante del efecto marginal del trabajo sobre la probabilidad de recibir un subsidio de AIS. En especial, $-3.5 * 100$ puntos porcentuales corresponde a la parte constante del efecto marginal de un trabajador más sobre la probabilidad de tener un subsidio de AIS.

$\hat{\beta}_4 = 1.0$ es la mitad de la parte variable del efecto marginal del trabajo sobre la probabilidad de recibir un subsidio de AIS. En especial, 100 puntos porcentuales corresponden a la mitad del aumento del efecto marginal de un trabajador más sobre la probabilidad de tener un subsidio de AIS.

En otras palabras, un trabajador más aumentará el efecto marginal del trabajo sobre la probabilidad de recibir un subsidio de AIS en $2 * 100$ puntos porcentuales.

Nota: El siguiente análisis no se esperaba. Ahora, si notan es de esperarse que el efecto marginal sea los trabajadores sea menor a -1 para 0 y 1 trabajadores y sea superior a 1 para más de 52 trabajadores (ver gráfico) por eso necesitaremos ajustar el efecto marginal, de tal manera como se ha presentado en el gráfico.

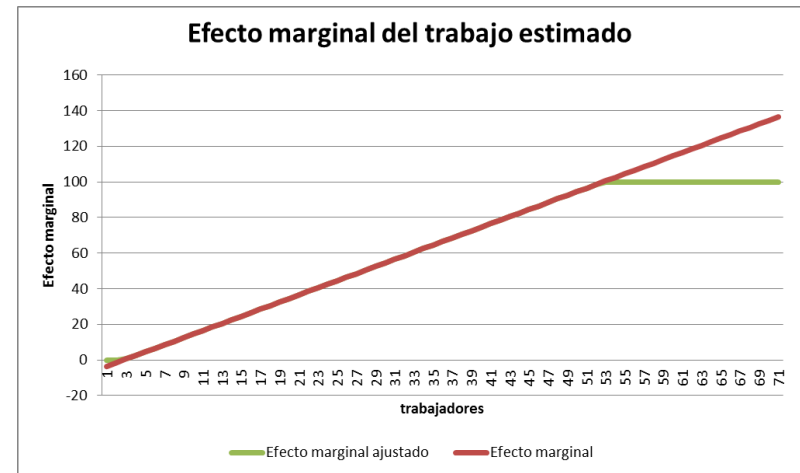


Tabla 1. Resultados de EasyReg.

Two-stage least squares:

Dependent variable:
Y = inf

X variables, including instrumental variables:
X(1) = open
X(2) = lnpcinc
X(3) = 1

WARNING: The effective degrees of freedom is only 12.
Therefore, the estimation results may be unreliable!

Endogenous X variables:
Y*(1)=inf

Exogenous X variables:
X*(1)=lnpcinc
X*(2)=1

2SLS estimation results for Y = inf

Parameters	Estimate	t-value (S.E.) [p-value]
Open	-0.3374	-26.693 (0.14166) [0.00000]
lnpcinc	0.3758	0.9650 (0.03775) [0.70000]
1	26.8993	0.37650 (0.03775) [0.82500]

[The p-values are two-sided and based on the normal approximation]

Standard error of the residuals = 1.2961
Residual sum of squares (RSS) = 60,64
Total sum of squares (TSS) = 606,4

R-squared = 0.81

Effective sample size (n) = 150
DW= 1.06