

Econometría 06216
Examen Final
Respuestas Sugeridas
Cali, Sábado 17 de Mayo de 2008

Profesores: Julio César Alonso
Geovanny Castro

Estudiante: _____
Código: _____

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **11** páginas; además, deben tener dos 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 esta 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. Al finalizar su examen entregue sus hojas de respuesta, así como las horas de preguntas.
8. 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 $y_t = AK_t^\alpha L_t^\beta$. Then ceteris paribus
 - a. α is the change in y per unit change in K
 - b. α is the percentage change in y per unit change in K
 - c. α is the percentage change in y per percentage change in K
 - d. α is none of the above because it is an elasticity
 - e. None of the above
2. Suppose your data produce the regression result $y = 10 + 3x$. Consider scaling the data to express them in a different base year dollar, by multiplying observations by 0.9. If both y and x are scaled, the new intercept and slope estimates will be
 - a. 10 and 3
 - b. 9 and 3
 - c. 10 and 2.7
 - d. 9 and 2.7
 - e. None of the above
3. Suppose you have estimated $wage = 5 + 3education + 2gender$, where gender is one for male and zero for female. If gender had been one for female and zero for male, this result would have been unchanged.
 - a. $wage = 5 + 3education - 2gender$
 - b. $wage = 7 + 3education + 2gender$
 - c. $wage = 7 + 3education - 2gender$
 - d. None of the above
4. You have estimated a logit model and found for a new individual that the estimated probability of her being a one (as opposed to a zero) is 40%. The benefit of correctly classifying this person is \$1,000, regardless of whether she is a one or a zero. The cost of

classifying this person as a one when she is actually a zero is \$500. What is the minimum the other misclassification can cost to make you classify her as a one?

- a. \$750
 - b. \$1,000
 - c. \$1250
 - d. \$1500
 - e. None of the above
5. If X and Y are random variables with $E[X] = 1$, $E[Y] = 2$, $var[X] = 2$, $var[Y] = 4$ and $cov(X,Y) = 2$, then $E[2X - 3Y]$ equals:
 - a. -4
 - b. 40
 - c. 8
 - d. -1
 - e. 10
 6. Si se presenta Asimetría a la derecha en la Distribución del término aleatorio de error de un modelo lineal con intercepto, entonces se puede afirmar que:
 - a. La Matriz de Varianzas-Covarianza de los Errores es Heteroscedastica.
 - b. Los EMCO son Inesgados.
 - c. Las Desviaciones Estándar de los EMCO son "Incorrectos"
 - d. Todas las Anteriores.
 - e. Ninguna de las anteriores.
 7. Cuando se presentan problemas de Heteroscedasticidad los Estimadores de Máxima Verosimilitud para los coeficientes asociados a pendientes y el intercepto son:
 - a. MELI.
 - b. Inesgados e Ineficientes.
 - c. sesgados e Ineficientes.
 - d. Lineales pero Sesgados.
 - e. Ninguna de las anteriores.
 8. Considere la siguiente ecuación de demanda de café estimada (error estándar entre paréntesis):

$$\ln \hat{Y}_t = 0.456 - 0.1832 \ln X_t$$

(0.2) (-0.3)

donde Y_t es la demanda de café en libras, teniendo en cuenta la significancia, y X_t es el

- precio real del café por libra. Teniendo en cuenta la significancia, es correcto afirmar que la demanda de café es:
- Inelástica al precio y negativa.
 - Perfectamente inelástica.
 - Elástica al precio y negativa.
 - No hay suficiente información para decidir.
 - Ninguna de las anteriores
9. El test de Breush-Pagan es un test que tiene una probabilidad de cometer error tipo I cuando concluimos que existe:
- Homoscedasticidad
 - Heteroscedasticidad
 - Autocorrelación
 - A y B.
 - Ninguna de las anteriores.
10. “Si hay un error en la medida de la variable dependiente del modelo, habrá problemas de sesgo”. La anterior afirmación es:
- Verdadera, porque los errores de medición distorsionan la información y la hacen inválida.
 - Verdadera, porque esto causará que se sobre-estime el valor real de los coeficientes.
 - Falsa.
 - A y B
11. A manufacturer has had to recall several models due to problems not discovered with its random final inspection procedures. This is an example of
- a type I error
 - a type II error
 - both types of error
 - neither type of error
 - None of the above
12. You have estimated a logit model to determine the probability that an individual is earning more than ten dollars an hour, with observations earning more than ten dollars an hour coded as ones; your estimated logit index function 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.
- Suppose you believe that the influence of experience depends on gender. To incorporate this into your logit estimation procedure 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
 - all of the above are possible solutions.
 - none of the above
13. A negative covariance between x and y means that whenever we obtain an x value that is greater than the mean of x
- we will obtain a corresponding y that is negative
 - we will obtain a corresponding y value smaller than the mean of y
 - the expected value of the corresponding y value will be negative
 - the expected value of the corresponding y value will be smaller than the mean of y
 - None of the above
14. Consider the two specifications
- $$y_t = \beta_0 + \beta_1 X_t^{-1} + \varepsilon_t \quad \text{and}$$
- $$y_t = \alpha X_t^{sen(\theta)} \varepsilon_t$$
- only the first specification can be estimated by a linear regression
 - only the second specification can be estimated by a linear regression
 - both specifications can be estimated by a linear regression
 - neither specification can be estimated by a linear regression
 - None of the above
15. Whenever the dependent variable is a fraction, using a linear functional form is OK if
- none of the dependent variable values are close to either zero or one

- b. only a few of the dependent variable values are close to either zero or one
- c. most of the dependent variable values are close to either zero or one
- d. it is never OK
- e. None of the above
16. El método de Durbin:
- a. Es un caso especial de los Mínimos Cuadrados Ordinarios (MCO).
- b. Es una manera de implementar los Mínimos cuadrados generalizados
- c. Es un caso especial de los Mínimos Cuadrados generalizados
- d. Asigna una menor influencia a las observaciones donde los datos presentan mayor ruido (varianza), y una mayor influencia a las observaciones donde los datos presentan menor ruido.
- e. B y c son ciertas
17. Si se emplea los estimadores de MCO en presencia de heteroscedasticidad, pero emplea cuidadosamente la “fórmula” sugerida por White para calcular la varianza de los parámetros, entonces:
- a. Se tendrá el mejor valor estimado que se pueda obtener, pues los MCO son MELI.
- b. Se aplicará después de haber agotado las posibilidades ofrecidas por los FGLS.
- c. Se tendrán valores estimados puntuales insesgados para los parámetros, pero errores estándares que no son tan pequeños como los que se obtienen por medio del método de mínimos cuadrados ordinarios.
- d. C y B
- e. Ninguna de las anteriores
18. Suponga el modelo $y_i = \beta_1 + \beta_1 x_i + \varepsilon_i$, en donde ε_i es el término de error, el cual es heteroscedástico con $\text{Var}(\varepsilon_i) = \alpha z_i^2$, donde z_i es una variable observable y α es un término constante desconocido. El modelo que debería ser utilizado para corregir el problema de heteroscedasticidad es:
- a. $y_i z_i = \beta_1 z_i + \beta_2 x_i z_i + \varepsilon_i^*$
- b. $(y_i/z_i) = \beta_1 (1/z_i) + \beta_2 (x_i/z_i) + \varepsilon_i^*$
- c. $y_i z_i^{1/2} = \beta_1 z_i^{1/2} + \beta_2 x_i z_i^{1/2} + \varepsilon_i^*$
- d. $(y_i/z_i^{1/2}) = \beta_1 (1/z_i^{1/2}) + \beta_2 (x_i/z_i^{1/2}) + \varepsilon_i^*$
- e. Ninguna de las anteriores
19. Una buena razón para emplear los estimadores de mínimos cuadrados en dos etapas es:
- a. que exista autocorrelación en el término aleatorio de error.
- b. SIEMPRE que la variable del lado izquierdo sea endógena.
- c. SIEMPRE que se estime una ecuación en el contexto de un sistema de ecuaciones simultáneas.
- d. SIEMPRE que se estime una ecuación de forma estructural
- e. Ninguna de las anteriores.
20. Bajo la hipótesis nula de no heteroscedasticidad, el valor del estadístico de la prueba de Goldfeld-Quandt se acerca a:
- a. Menos de cero
- b. 0
- c. -1
- d. 0.5
- e. Ninguna de las anteriores.
21. In the case of the simple regression model $Y_i = \beta_0 + \beta_1 X_i + u_i$, $i = 1, \dots, n$, when X and u are correlated, then
- a. the OLS estimator is biased in small samples only.
- b. OLS and 2SLS produce the same estimate.
- c. X is exogenous.
- d. the OLS estimator is inconsistent.
- e. None of the above
22. The reason why estimators have a sampling distribution is that
- a. economics is not a precise science.
- b. individuals respond differently to incentives.

- c. in real life you typically get to sample many times.
 d. the values of the explanatory variable and the error term differ across samples.
 e. None of the above
23. Consider the following two statements:
 If you reject a null using a one-tailed test, then you will also reject it using a two-tailed test at the same significance level;
 For a given level of significance, the critical value of t gets closer to zero as the sample size increases.
 a. both statements are true
 b. neither statement is true
 c. only the first statement is true
 d. only the second statement is true
 e. None of the above
24. Adding an irrelevant explanatory variable which is orthogonal to the other explanatory variables causes
 a. bias and no change in variance
 b. bias and an increase in variance
 c. no bias and no change in variance
 d. no bias and an increase in variance
 e. None of the above
25. Suppose you regress y on x and the square of x .
 a. Estimates will be biased with large variances
 b. It doesn't make sense to use the square of x as a regressor
 c. The regression will not run because these two regressors are perfectly correlated
 d. There should be no problem with this.
 e. None of the above
26. You have regressed $\ln(\text{wage})$ on Ed (years of education). The slope and intercept coefficient estimates are random variables, so they have a covariance. A positive covariance here implies that
 a. the slope coefficient is positive
 b. high values of wage correspond to high values of Ed
 c. low values of the intercept estimate correspond to low values of the slope estimate
 d. all of the above
 e. None of the above
27. Suppose $\text{wage}_i = \alpha + \beta x_i + \varepsilon_i$ and we have 100 observations on wage and x , with average values 70 and 30, respectively. We have run a regression to estimate the slope of x as 2.0. Consider now a new individual whose age is 20. For this individual the predicted wage from this regression is
 a. 50
 b. 60
 c. impossible to predict without knowing the intercept estimate
 d. None of the above
28. Suppose you have run the following regression:

$$y_i = \alpha + \beta x_i + \chi \text{Urban}_i + \delta \text{Immigrant}_i + \phi \text{Urban}_i * \text{Immigrant}_i + \varepsilon_i$$
 where Urban is a dummy indicating that an individual lives in a city ($\text{Urban}=1$) rather than in a rural area ($\text{Urban}=0$), and Immigrant is a dummy indicating that an individual is an immigrant ($\text{Immigrant}=1$) rather than a native. The coefficient δ is interpreted as the ceteris paribus difference in y between:
 a. an immigrant and a native
 b. a rural immigrant and a rural native
 c. an urban immigrant and an urban native
 d. an urban immigrant and a rural native
 e. None of the above
29. If the expected value of the error term is 5, then after running an OLS regression of y on the relevant explanatory variables and a constant term
 a. the average of the residuals should be approximately 5
 b. the average of the residuals should be exactly zero
 c. the average of the residuals should be exactly five
 d. nothing can be said about the average of the residuals
 e. None of the above
30. In a logit regression, to report the influence of an explanatory variable x on

- the probability of observing a one for the dependent variable we report
- the slope coefficient estimate for x
 - the average of the slope coefficient estimates for x of all the observations in the sample
 - the slope coefficient estimate for x for the average observation in the sample
 - none of these
 - all of these could be a good way to report.
31. To find the maximum likelihood estimates the computer searches over all possible values of the
- dependent variable
 - independent variables
 - coefficients
 - the independent variables and the coefficients
 - None of the above
32. Consider the multivariate classical linear regression model: $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ where \mathbf{Y} is the $(n \times 1)$ matrix of dependent variables, \mathbf{X} is the $(n \times k)$ matrix of explanatory variables, β is the $(k \times 1)$ matrix of unknown coefficients and ε is the $(n \times 1)$ matrix of the random effect on \mathbf{Y} . Which of the following statements is correct?
- Given that \mathbf{C}' is a $(n \times 1)$ vector of constants $[c_1, c_2, \dots, c_n]$, the quantity $\mathbf{C}(\mathbf{Y}\mathbf{Y}')\mathbf{C}'$ is a *quadratic form* in \mathbf{Y} .
 - $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Y}'\mathbf{X}$ is an unbiased estimator of β .
 - The assumption of *Homoskedasticity alone* implies that the variance-covariance matrix of ε is given by $\sigma^2\mathbf{I}_n$ where \mathbf{I}_n is the $(n \times n)$ identity matrix and σ^2 is unknown.
 - The estimator $\hat{\beta}_1$ for the parameter β_1 is a *quadratic form* in the \mathbf{Y} values.
 - None of the above
33. Consider the classical linear regression model: $Y_i = \beta_1 + \beta_2 X_{2i} + \varepsilon_i$
- In the following set of questions, the parameters β_1 and β_2 are estimated by Ordinary Least Squares (OLS) method to generate estimators $\hat{\beta}_1$ and $\hat{\beta}_2$. Which of the following is a correct statement?
- $\hat{\beta}_2$ will not be an efficient estimator of β_2 if $\text{Cov}(\varepsilon_i, \varepsilon_j)$ is not zero for some i and j ($i \neq j$).
 - The covariance of $\hat{\beta}_1$ and $\hat{\beta}_2$ is positive if the mean of the X_{2i} is zero.
 - Suppose that $\text{Cov}(X_{2i}, \varepsilon_i) > 0$. The OLS method will produce unbiased estimators provided that the sum of the error terms e_i equals 0.
 - We need to assume that the random term ε_i has a normal distribution in order to show that $\hat{\beta}_1$ and $\hat{\beta}_2$ are consistent estimators of β_1 and β_2 respectively.
 - None of the above
34. In the binary dependent variable model, a predicted value of 0.6 means that:
- the most likely value the dependent variable will take on is 60 percent.
 - given the values for the explanatory variables, there is a 60 percent probability that the dependent variable will equal one.
 - the model makes little sense, since the dependent variable can only be 0 or 1.
 - given the values for the explanatory variables, there is a 40 percent probability that the dependent variable will equal one.
 - None of the above
35. 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. $\hat{\beta}$ is an unambiguously better estimator than $\tilde{\beta}$.
- b. $\tilde{\beta}$ is an unambiguously better estimator than $\hat{\beta}$.
- c. $\hat{\beta}$ has to be the OLS estimator.
- d. a. an c. are true.
- e. None of the above
36. ¿Cuál de las siguientes pruebas para comprobar autocorrelación es válida aun cuando los regresores en la ecuación original incluyen un rezago de la variable dependiente, por ejemplo, en la ecuación: $y_t = \beta_1 + \beta_2 x_t + \beta_3 y_{t-1} + \epsilon_t$?
- a. Una prueba t de una regresión hecha por MCO entre la variable dependiente y su variable rezaga sin incluir intercepto.
- b. La Prueba de Durbin-Watson
- c. La Prueba de Breush-Pagan.
- d. La prueba de White
- e. Ninguna de las anteriores
37. Los estimadores MCO **NO** serán sesgados si:
- a. Cualquiera de las variables independientes es medida con error.
- b. Una de las variables binarias es medida con error.
- c. Se omite una variable independiente que no se encuentra correlacionada con las variables independientes incluidas en la ecuación.
- d. Todas las anteriores
- e. Ninguna de las anteriores
38. Si el término de error está autocorrelacionado, entonces los coeficientes estimados por el método de MCO son:
- a. Insesgados y eficientes
- b. Insesgados e ineficientes.
- c. Sólo ineficientes.
- d. Sólo insesgados.
- e. Ninguna de las anteriores
39. El problema de simultaneidad implica:
- a. La posible inexistencia de un sesgo en los estimadores MCO.
- b. La necesidad de emplear los estimadores de MCG.
- c. La existencia de una variable (explicativa o dependiente) relacionada con el error.
- d. Todas las anteriores
- e. Ninguna de las anteriores
40. Ante la presencia de un problema de multicolinealidad no perfecta:
- a. La mejor forma de solucionarla es eliminando una de las variables correlacionadas.
- b. En ningún caso es posible solucionarlas.
- c. Se considera que los betas están correlacionados con los otros betas y por eso uno solo de ellos recoge todo el efecto de las variables correlacionadas.
- d. Todas las anteriores
- e. Ninguna de las anteriores.
41. El test de Durbin y Watson es un test de:
- a. No autocorrelación
- b. Autocorrelación
- c. Heteroscedasticidad
- d. A y B.
- e. Ninguna de las anteriores.
42. En un modelo de regresión lineal con una tendencia lineal (time trend), el coeficiente asociado a la tendencia puede ser:
- a. un estimador de la tasa de crecimiento % de la variable dependiente (ceteris paribus).
- b. El cambio en otra de las variables independientes por periodo de tiempo (ceteris paribus).
- c. (a) y (b) pueden ser correctas.
- d. Ninguna de las anteriores.
43. Comparando el método de Máxima Verosimilitud (MV) y el de Mínimos Cuadrados Ordinarios para la estimación de un modelo con variable dicotómica como dependiente, NO se puede afirmar que:
- a. MCO escoge la línea que mejor se ajusta por medio de la minimización de la suma de las desviaciones (en

- sentido vertical) de cada uno de los valores de Y con respecto a esta línea.
- b. MV escoge la línea que mejor se ajusta maximizando la probabilidad conjunta de observar las n-observaciones independientes sobre Y en la muestra.
- c. Tendrán resultados diferentes.
- d. Todas las anteriores
- e. Ninguna de las anteriores.
44. La autocorrelación en los datos es un problema porque:
- a. El método de MCO asume que los datos no son correlacionados y calcula los estimadores puntuales de los betas de la regresión de manera acorde.
- b. Sesga el estimador de la matriz de varianzas y covarianzas.
- c. (a) y (b) son correctas.
- d. Ninguna de las anteriores
45. Qué supuesto se requiere para que los estimadores de MCO sean insesgados?
- a. Correr el modelo por MCO es suficiente, ya que el teorema de Gauss-Markov afirma que cualquier especificación de un modelo por MCO dará como resultado estimadores insesgados.
- b. Homoscedasticidad: $\text{Var}(\varepsilon_i) = \sigma^2$
- c. No autocorrelación: $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ para $i \neq j$.
- d. Ninguna de las anteriores
46. La heteroscedasticidad es un problema que puede presentarse en:
- a. Los modelos que emplean datos de corte transversal.
- b. Los modelos que emplean datos de series de tiempo.
- c. A y B
- d. Ninguna de las anteriores
47. La diferencia entre las pruebas de autocorrelación de Box-Pierce y Durbin-Watson es:
- a. La primera detecta solamente autocorrelación de primer orden.
- b. La segunda detecta solamente autocorrelación de primer orden.
- c. Ambas detectan solamente autocorrelación de primer orden, por lo tanto, éste no sería un factor diferenciador.
- d. Ninguna de las anteriores.
48. Los estimadores de máxima verosimilitud para regresiones lineales, dan como resultado:
- a. Estimaciones insesgadas de los betas. Ninguna de las anteriores
- b. Estimaciones eficientes de la matriz de varianzas y covarianzas. Ninguna de las anteriores
- c. Estimaciones insesgadas tanto de los betas como de la matriz de varianzas y covarianzas. Ninguna de las anteriores
- d. Estimaciones sesgadas tanto de los betas como de la matriz de varianzas y covarianzas. Ninguna de las anteriores
49. La diferencia entre 8% y 5% es:
- a. 3 puntos porcentuales
- b. 3 por ciento
- c. Es indistinto a o b es lo mismo.
- d. Dependiendo del contexto en el que se comparen.
- e. Ninguna de las anteriores
50. Uno de los siguientes es un supuesto de Gauss-Markov y puede ser relajado y permitir que los estimadores continúen siendo MELI.
- a. Omisión de variables explicativas no relevantes.
- b. X estocásticas.
- c. Valor esperado del error igual a 0.
- d. X no estocásticas.
- e. Ninguna de las anteriores

II. (20 puntos)

Un investigador está interesado en estimar un modelo tipo IS-LM. Usted acaba de ser contratado como asistente de la investigación en reemplazo del asistente que se ira a estudiar al exterior. El anterior asistente estimó una de las ecuaciones simultáneas que conforman el sistema. Estos resultados se reporta al final del examen.

En donde R_t , M_t , Y_t , e Inv_t , corresponden al tipo de interés, la oferta monetaria (en millones de pesos constates de 1994), el PIB (en millones de pesos constates de 1994) y la inversión (en millones de pesos constates de 1994) para el año t, respectivamente. Así mismo, D_t corresponde a una variable dummy que toma el valor de uno para años en los que el Banco Central ha sido independiente de la rama ejecutiva. Para ello dispone de los datos anuales para el periodo 1966-2006 de este país. Teniendo en cuenta esta información, responda:

- a) De acuerdo a los cálculos del econometrista, determine cuáles son las variables endógenas y cuáles las exógenas del sistema. **(5 puntos)**
- b) Es posible afirmar que este modelo estimado presenta síntomas de autocorrelación negativa? Explique su respuesta de la manera más clara posible. **(5 puntos)**
- c) Teniendo en cuenta la significancia y omitiendo el problema encontrado (o no), interprete los coeficientes estimados **(5 puntos)**
- d) Explique claramente como podría comprobar la hipótesis de que la independencia del Banco de la república ha sido una buena idea para esta economía. **(5 puntos)**

III. (30 puntos)

Un investigador está interesado en estimar un modelo que permita determinar las variables que afectan el hecho de que un individuo sea jugador de chance (J_i es una variable dummy que toma el valor de uno si el individuo es jugador de chance y cero en caso contrario) y el monto de dinero que los individuos apuestan en cualquier juego de azar en el Valle del Cauca (A_i representa el gasto mensual en juegos de azar del individuo i en pesos para el juego de chance).

Para lograr su objetivo, el investigador emplea una base de datos que corresponde a encuestas realizadas a una muestra representativa de jugadores en el Valle del Cauca. Así mismo, emplea las dos siguientes especificaciones:

$$\ln(A_i) = \alpha_1 + \alpha_2 \ln(Y_i) + \alpha_3 G_i + \alpha_4 \ln(edu_i) + \alpha_5 J_i + \varepsilon_i \tag{1}$$

$$J_i = \beta_1 + \beta_2 G_i + \beta_3 \ln(edu_i) + \mu_i \tag{2}$$

donde Y_i , G_i , y edu_i corresponden a al ingreso del individuo i (medio en miles de pesos por mes), una variable dummy que toma el valor de uno si el individuo i es un hombre y cero en caso contrario y los años de educación del individuo i, respectivamente.

- a) Rápidamente, discuta que tipo de problema(s) econométrico(s) puede esperar (a priori) el investigador tener si estima (1) y (2) por MCO y si es del caso sugiera un método alternativo de estimación. **Sea lo más concreto y preciso posible(10 puntos)**

Ignorando los problemas (o no) que se pueden presentar en la especificación (2), se recoge la información necesaria para estimarla por medio del método de MCO. Para esto el asisten del investigador preparó las siguientes matrices:

$$X^T X = \begin{bmatrix} 2500 & 500 & 0 \\ 500 & 200 & 0 \\ 0 & 0 & 12 \end{bmatrix} \qquad X^T y = \begin{bmatrix} 300 \\ 200 \\ 180 \end{bmatrix}$$

- b) De acuerdo a estos datos, determine (si es posible) cuál es el tamaño de la muestra empleada, cuál es la proporción de NO jugadores de chance según la muestra y que proporción de la muestra corresponde a mujeres. **(5 puntos)**
- c) Encuentre los estimadores MCO del modelo (2). **(10 Puntos)**
- d) Interprete el significado de cada uno de los coeficientes estimados. **(5 Puntos)**

Resultados de EasyReg.

Two-stage least squares:
 Dependent variable:
 $Y = \ln(Y)$
 X variables, including instrumental variables:
 $X(1) = \ln(INV)$
 $X(2) = \ln(M)$
 $X(3) = \ln(R)$
 $X(4) = \text{LAG1}[\ln(Y)]$
 $X(5) = \ln(\text{LAG1}[R])$
 $X(6) = D * \ln(INV)$
 $X(7) = D * \ln(M)$
 $X(8) = D * \ln(\text{LAG1}[R])$
 $X(9) = D * (\text{LAG1}[\ln(Y)])$
 $X(10) = 1$
 Endogenous X variable:
 $Y^*(1) = \ln(R)$
 Exogenous X variables:
 $X^*(1) = \ln(INV)$
 $X^*(2) = D * \ln(INV)$
 $X^*(2) = 1$
 2SLS estimation results for $Y = Y$

Variables	2SLS estimate	t-value
		[p-value]
$\ln(INV)$	18.6928	5.339
		[0,00000]
$\ln(R)$	-1.30960	-3.903
		[0,00009]
$D * \ln(INV)$	3.0960	4.903
		[0,1009]
1	34.99671	6.118
		[0,00000]

[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-square = 0.80
 Effective sample size (n) = 40
 DW = 0.5

HOJA DE RESPUESTAS PREGUNTAS DE SELECCIÓN MÚLTIPLE

NOMBRE: _____

En está hoja marque la respuesta correcta.

- | | A | B | C | D | E | | A | B | C | D | E |
|-----|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 26. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 27. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 28. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 29. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 30. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 31. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 32. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 33. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 34. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 35. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 36. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 37. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 13. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 38. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 14. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 39. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 40. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 41. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 42. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 43. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 19. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 44. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 20. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 45. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 21. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 46. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 22. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 47. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 23. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 48. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 24. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 49. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 25. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | 50. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

MAS sin reposición

$$\frac{n}{N} \quad \frac{N!}{n!(N-n)!}$$

$$S_{\bar{y}} = \sqrt{\left(1 - \frac{n}{N}\right)} \frac{S}{\sqrt{n}}$$

$$\hat{P} = \frac{\sum_{i=1}^n X_i}{n} \quad S_{\hat{P}}^2 = \frac{N-n}{N-1} \frac{\hat{P}\hat{Q}}{n}$$

MAS con reposición

$$1 - \left(\frac{N-1}{N}\right)^n \quad N^n$$

$$S_{\bar{y}} = \frac{S}{\sqrt{n}}$$

$$S_{\hat{P}}^2 = \frac{\hat{P}\hat{Q}}{n}$$

MEA

$$W_h = \frac{N_h}{N} \text{ para } h = 1, 2, \dots, H$$

$$n = \frac{\frac{n!(N-n)!}{N!} \frac{z_{\alpha/2}^2 S^2}{\delta^2}}{1 + \frac{1}{N} \left(\frac{z_{\alpha/2}^2 S^2}{\delta^2} \right)}$$

$$n_{(p)} = \frac{\frac{z_{\alpha/2}^2 \hat{P}\hat{Q}}{\delta^2}}{\frac{N-1}{N} + \frac{1}{N} \left(\frac{z_{\alpha/2}^2 \hat{P}\hat{Q}}{\delta^2} \right)}$$

$$\frac{1}{N^n}$$

$$n_0 = \frac{z_{\alpha}^2 S^2}{\delta^2}$$

$$n_{(p)} = \frac{z_{\alpha/2}^2 \hat{P}\hat{Q}}{\delta^2}$$

$$n = \sum_{h=1}^H n_h$$

$$\bar{y} = \sum_{h=1}^H W_h \bar{y}_h \quad \bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{h,i}}{n_h}$$

$$S_{\bar{y}_h} = \sqrt{\left(1 - \frac{n_h}{N_h}\right)} \frac{S_h}{\sqrt{n_h}}$$

$$n_h = n \frac{W_h S_H}{\sum_{h=1}^H W_h S_H}$$

Muestreo por Conglomerado

$$\bar{y}_{congl} = \frac{\sum_{i=1}^{n_c} \sum_{j=1}^M y_{i,j}}{n_c}$$

$$S_{congl}^2 = \frac{\sum_{i=1}^{n_c} \left(\left(\sum_{j=1}^M y_{i,j} \right) - \bar{y}_{congl} \right)^2}{n_c - 1}$$

$$S_{\bar{y}_{congl}}^2 = \frac{N_C (N_C - n_C)}{n_C} S_{congl}^2$$

$$Var[\bar{y}] = \sum_{h=1}^H W_h^2 Var[\bar{y}_h]$$

$$n = \frac{\frac{z_{\alpha/2}^2 \left(\sum_{h=1}^H W_h S_H \right)^2}{\delta^2}}{1 + \frac{1}{N} \left(\frac{z_{\alpha/2}^2 \left(\sum_{h=1}^H W_h S_H \right)^2}{\delta^2} \right)}$$

$$\bar{y} = \frac{\sum_{i=1}^{n_c} \sum_{j=1}^M y_{i,j}}{n} = \frac{\bar{y}_{congl}}{M}$$

$$S^2 = \frac{\sum_{i=1}^{n_c} \sum_{j=1}^M (y_{i,j} - \bar{y})^2}{n - 1}$$

$$S_{\bar{y}}^2 = \frac{(N_C - n_C) S_{congl}^2}{N_C n_C M^2}$$

Prof: Julio César Alonso C

$$n_c = \frac{\frac{z_{\alpha/2}^2 S_{congl}^2}{\delta^2 M^2}}{1 + \frac{1}{N_c} \left(\frac{z_{\alpha/2}^2 S_{congl}^2}{\delta^2 M^2} \right)} n = n_c M$$

$$X^T X = \begin{bmatrix} n & \sum_{i=1}^n X_{1i} & \sum_{i=1}^n X_{2i} & \cdots & \sum_{i=1}^n X_{ki} \\ \sum_{i=1}^n X_{1i}^2 & \sum_{i=1}^n X_{1i} X_{2i} & \cdots & \sum_{i=1}^n X_{1i} X_{ki} \\ & \sum_{i=1}^n X_{2i}^2 & \ddots & \sum_{i=1}^n X_{2i} X_{ki} \\ & & \ddots & \vdots \\ & & & \sum_{i=1}^n X_{ki}^2 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i X_{1i} \\ \sum_{i=1}^n y_i X_{2i} \\ \vdots \\ \sum_{i=1}^n y_i X_{ki} \end{bmatrix} \quad y^T y = \sum_{i=1}^n y_i^2$$

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 = y^T y - n\bar{y}^2 \quad SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad s^2 = \frac{SSE}{n-k} = \frac{y^T y - \hat{\beta}^T X^T y}{n-k}$$

$$Var[\hat{\beta}] = \sigma^2 (X^T X)^{-1}$$

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \hat{\beta}^T X^T y - n\bar{y}^2 \quad t = \frac{\hat{\beta}_i - c}{s_{\hat{\beta}_i}}$$

$$F_c = \frac{(c - R\hat{\beta})^T (R(X^T X)^{-1} R^T)^{-1} (c - R\hat{\beta}) / r}{SSE / (n-k)}$$

$$F_C = \frac{R^2 / (k-1)}{(1-R^2) / (n-k)} = \frac{MSR}{MSE}$$

$$F_c = \frac{(SSE_R - SSE_U) / r}{SSE_U / (n-k)} \quad R^2 = \frac{SSR}{SST}$$

$$\hat{\beta}_i \pm t_{\frac{\alpha}{2}, n-k} s_{\hat{\beta}_i} \quad \bar{R}^2 = 1 - (1-R^2) \frac{n-1}{n-k}$$

$$\hat{y}_p = x_p^T \hat{\beta} \quad x_p^T = (1 \quad x_{1p} \quad x_{2p} \quad \cdots \quad x_{kp})$$

$$\hat{y}_p \pm t_{\frac{\alpha}{2}, n-k} \sqrt{\sigma^2 x_p^T (X^T X)^{-1} x_p}$$

$$\hat{y}_p \pm t_{\frac{\alpha}{2}, n-k} \sqrt{\sigma^2 \left[1 + x_p^T (X^T X)^{-1} x_p \right]}$$

Prof: Julio César Alonso C

$$\hat{\beta}_j^E = \hat{\beta}_j \frac{s_{x_j}}{s_y}, \quad j = 2, 3, \dots, k$$

$$E_j = \hat{\beta}_j \frac{\bar{X}_j}{\bar{y}}$$

$$h = \hat{\rho} \sqrt{\frac{n}{1 - n(\overline{Var(\hat{\alpha})})}} \quad \text{donde}$$

$$y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + \alpha y_{t-1} + \varepsilon_t$$

Test de Heteroscedasticidad

Goldfeld y Quand: $F_{GQ} = \frac{SSE_2}{SSE_1} \sim F_{(n-d-2k, n-d-2k)}$

Breush-Pagan: $\frac{\hat{\varepsilon}_i^2}{\hat{\sigma}^2} = \gamma + \delta Z_i + \mu_i, \quad BP = \frac{SSR}{2} \sim \chi_g^2$

White: $\hat{\varepsilon}_i^2 = \gamma + \sum_{m=1}^k \sum_{j=1}^k \delta_s X_{mi} X_{ji} + \mu_i, \quad W_a = nR^2 \sim \chi_g^2$

Cantidades Importantes

$$\sqrt{2} = 1.414$$

$$\sqrt{3} = 1.732$$

$$\sqrt{5} = 2.236$$

$$\sqrt{10} = 3.162$$

$$\sqrt{13} = 3.606$$

Test de Autocorrelación

Durbin-Watson $DW \approx 2(1 - \hat{\rho})$

Ho	Sí	Decisión
$H_0 : \rho = 0$	$d_u < DW < 4 - d_u$	A
No auto +	$0 < DW < d_l$	R
No auto -	$4 - d_l < DW < 4$	R

Área de indecisión $d_l < DW < d_u$ y $4 - d_u < DW < 4 - d_l$

d_l y d_u para el test de DW al nivel de significancia del 5%

N	k-1=1		k-1=2		k-1=3	
	d _l	d _u	d _l	d _u	d _l	d _u
50	1.50	1.59	1.46	1.63	1.42	1.67
60	1.55	1.62	1.51	1.65	1.48	1.69
95	1.64	1.69	1.62	1.71	1.60	1.73
100	1.65	1.69	1.63	1.72	1.61	1.74

Condición de Orden

$k_i > g_i - 1$ *sobre-identificada*

$k_i = g_i - 1$ *perfectamente identificada*

Econometría 06216
Examen Final
Respuestas Sugeridas
Cali, Sábado 17 de Mayo de 2008

Profesores: Julio César Alonso
Geovanny Castro

Estudiante: _____
Código: _____

Instrucciones:

1. Lea cuidadosamente todas las preguntas e instrucciones.
2. Este examen consta de **5** páginas; además, deben tener una hoja 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 esta 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. Al finalizar su examen entregue sus hojas de respuesta, así como las hojas de preguntas.
8. 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 $y_t = AK_t^\alpha L_t^\beta$. Then ceteris paribus
 - a. α is the change in y per unit change in K
 - b. α is the percentage change in y per unit change in K
 - c. α is the percentage change in y per percentage change in K
 - d. α is none of the above because it is an elasticity
 - e. None of the above
2. Suppose your data produce the regression result $y = 10 + 3x$. Consider scaling the data to express them in a different base year dollar, by multiplying observations by 0.9. If both y and x are scaled, the new intercept and slope estimates will be
 - a. 10 and 3
 - b. 9 and 3
 - c. 10 and 2.7
 - d. 9 and 2.7
 - e. None of the above
3. Suppose you have estimated $wage = 5 + 3education + 2gender$, where gender is one for male and zero for female. If gender had been one for female and zero for male,
 - a. this result would have been unchanged.
 - b. $wage = 5 + 3education - 2gender$
 - c. $wage = 7 + 3education + 2gender$
 - d. $wage = 7 + 3education - 2gender$
 - e. None of the above
 - f. Ojo en el examen había un error. La primera opción no aparecía numerada y por esto originalmente la respuesta correcta era la opción c. En el examen se calificará como respuesta

correcta la opción c ($wage = 7 + 3education - 2gender$)

4. You have estimated a logit model and found for a new individual that the estimated probability of her being a one (as opposed to a zero) is 40%. The benefit of correctly classifying this person is \$1,000, regardless of whether she is a one or a zero. The cost of classifying this person as a one when she is actually a zero is \$500. What is the minimum the other misclassification can cost to make you classify her as a one?
 - a. \$750
 - b. \$1,000
 - c. \$1250
 - d. \$1500
 - e. None of the above
5. If X and Y are random variables with $E[X] = 1$, $E[Y] = 2$, $var[X] = 2$, $var[Y] = 4$ and $cov(X,Y) = 2$, then $E[2X - 3Y]$ equals:
 - a. -4
 - b. 40
 - c. 8
 - d. -1
 - e. 10
6. Si se presenta Asimetría a la derecha en la Distribución del término aleatorio de error de un modelo lineal con intercepto, entonces se puede afirmar que:
 - a. La Matriz de Varianzas-Covarianza de los Errores es Heteroscedastica.
 - b. Los EMCO son Insesgados.
 - c. Las Desviaciones Estándar de los EMCO son "Incorrectos"
 - d. Todas las Anteriores.
 - e. Ninguna de las anteriores.
7. Cuando se presentan problemas de Heteroscedasticidad los Estimadores de Máxima Verosimilitud para los coeficientes asociados a pendientes y el intercepto son:
 - a. MELI.
 - b. Insesgados e Ineficientes.
 - c. sesgados e Ineficientes.
 - d. Lineales pero Sesgados.

- e. Ninguna de las anteriores.
8. Considere la siguiente ecuación de demanda de café estimada (error estándar entre paréntesis):
- $$\ln \hat{Y}_t = 0.456 - 0.1832 \ln X_t, \text{ donde } Y_t \text{ es la demanda de café en libras, teniendo en cuenta la significancia, y } X_t \text{ es el precio real del café por libra. Teniendo en cuenta la significancia, es correcto afirmar que la demanda de café es:}$$
- (0.2) (-0.3)
- a. Inelástica al precio y negativa.
- b. **Perfectamente inelástica.**
- c. Elástica al precio y negativa.
- d. No hay suficiente información para decidir.
- e. Ninguna de las anteriores
9. El test de Breush-Pagan es un test que tiene una probabilidad de cometer error tipo I cuando concluimos que existe:
- a. Homoscedasticidad
- b. **Heteroscedasticidad**
- c. Autocorrelación
- d. A y B.
- e. Ninguna de las anteriores.
10. “Si hay un error en la medida de la variable dependiente del modelo, habrá problemas de sesgo”. La anterior afirmación es:
- a. Verdadera, porque los errores de medición distorsionan la información y la hacen inválida.
- b. Verdadera, porque esto causará que se sobre-estime el valor real de los coeficientes.
- c. **Falsa.**
- d. A y B
11. A manufacturer has had to recall several models due to problems not discovered with its random final inspection procedures. This is an example of
- a. a type I error
- b. **a type II error**
- c. both types of error
- d. neither type of error
- e. None of the above
12. You have estimated a logit model to determine the probability that an individual is earning more than ten dollars an hour, with observations earning more than ten dollars an hour coded as ones; your estimated logit index function 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.
- Suppose you believe that the influence of experience depends on gender. To incorporate this into your logit estimation procedure 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. all of the above are possible solutions.
- e. none of the above
13. A negative covariance between x and y means that whenever we obtain an x value that is greater than the mean of x
- a. we will obtain a corresponding y that is negative
- b. we will obtain a corresponding y value smaller than the mean of y
- c. the expected value of the corresponding y value will be negative
- d. **the expected value of the corresponding y value will be smaller than the mean of y**
- e. None of the above
14. Consider the two specifications
- $$y_t = \beta_0 + \beta_1 X_t^{-1} + \varepsilon_t \quad \text{and}$$
- $$y_t = \alpha X_t^{\text{sen}(\theta)} \varepsilon_t$$
- a. only the first specification can be estimated by a linear regression

- b. only the second specification can be estimated by a linear regression
- c. **both specifications can be estimated by a linear regression**
- d. neither specification can be estimated by a linear regression
- e. None of the above
15. Whenever the dependent variable is a fraction, using a linear functional form is OK if
- a. **none of the dependent variable values are close to either zero or one**
- b. only a few of the dependent variable values are close to either zero or one
- c. most of the dependent variable values are close to either zero or one
- d. it is never OK
- e. None of the above
16. El método de Durbin:
- a. Es un caso especial de los Mínimos Cuadrados Ordinarios (MCO).
- b. **Es una manera de implementar los Mínimos cuadrados generalizados**
- c. Es un caso especial de los Mínimos Cuadrados generalizados
- d. Asigna una menor influencia a las observaciones donde los datos presentan mayor ruido (varianza), y una mayor influencia a las observaciones donde los datos presentan menor ruido.
- e. B y c son ciertas
17. Si se emplea los estimadores de MCO en presencia de heteroscedasticidad, pero emplea cuidadosamente la "fórmula" sugerida por White para calcular la varianza de los parámetros, entonces:
- a. Se tendrá el mejor valor estimado que se pueda obtener, pues los MCO son MELI.
- b. Se aplicará después de haber agotado las posibilidades ofrecidas por los FGLS.
- c. Se tendrán valores estimados puntuales insesgados para los parámetros, pero errores estándares que no son tan pequeños como los que se obtienen por medio del método de mínimos cuadrados ordinarios.
- d. **C y B**
- e. Ninguna de las anteriores
18. Suponga el modelo $y_i = \beta_1 + \beta_2 x_i + \epsilon_i$, en donde ϵ_i es el término de error, el cual es heteroscedástico con $\text{Var}(\epsilon_i) = \alpha z_i^2$, donde z_i es una variable observable y α es un término constante desconocido. El modelo que debería ser utilizado para corregir el problema de heteroscedasticidad es:
- a. $y_i z_i = \beta_1 z_i + \beta_2 x_i z_i + \epsilon_i^*$
- b. **$(y_i/z_i) = \beta_1 (1/z_i) + \beta_2 (x_i/z_i) + \epsilon_i^*$**
- c. $y_i z_i^{1/2} = \beta_1 z_i^{1/2} + \beta_2 x_i z_i^{1/2} + \epsilon_i^*$
- d. $(y_i/z_i^{1/2}) = \beta_1 (1/z_i^{1/2}) + \beta_2 (x_i/z_i^{1/2}) + \epsilon_i^*$
- e. Ninguna de las anteriores
19. Una buena razón para emplear los estimadores de mínimos cuadrados en dos etapas es:
- a. que exista autocorrelación en el término aleatorio de error.
- b. SIEMPRE que la variable del lado izquierdo sea endógena.
- c. SIEMPRE que se estime una ecuación en el contexto de un sistema de ecuaciones simultáneas.
- d. SIEMPRE que se estime una ecuación de forma estructural
- e. **Ninguna de las anteriores.**
20. Bajo la hipótesis nula de no heteroscedasticidad, el valor del estadístico de la prueba de Goldfeld-Quandt se acerca a:
- a. Menos de cero
- b. 0
- c. -1
- d. 0.5
- e. Ninguna de las anteriores.
21. In the case of the simple regression model $Y_i = \beta_0 + \beta_1 X_i + u_i$, $i = 1, \dots, n$, when X and u are correlated, then

- a. the OLS estimator is biased in small samples only.
- b. OLS and 2SLS produce the same estimate.
- c. X is exogenous.
- d. **the OLS estimator is inconsistent.**
- e. None of the above
22. The reason why estimators have a sampling distribution is that
- a. economics is not a precise science.
- b. individuals respond differently to incentives.
- c. in real life you typically get to sample many times.
- d. **the values of the explanatory variable and the error term differ across samples.**
- e. None of the above
23. Consider the following two statements:
If you reject a null using a one-tailed test, then you will also reject it using a two-tailed test at the same significance level;
For a given level of significance, the critical value of t gets closer to zero as the sample size increases.
- a. both statements are true
- b. neither statement is true
- c. only the first statement is true
- d. **only the second statement is true**
- e. None of the above
24. Adding an irrelevant explanatory variable which is orthogonal to the other explanatory variables causes
- a. bias and no change in variance
- b. bias and an increase in variance
- c. **no bias and no change in variance**
- d. no bias and an increase in variance
- e. None of the above
25. Suppose you regress y on x and the square of x .
- a. Estimates will be biased with large variances
- b. It doesn't make sense to use the square of x as a regressor
- c. The regression will not run because these two regressors are perfectly correlated
- d. **There should be no problem with this.**
- e. None of the above
26. You have regressed $\ln(\text{wage})$ on Ed (years of education). The slope and intercept coefficient estimates are random variables, so they have a covariance. A positive covariance here implies that
- a. the slope coefficient is positive
- b. high values of wage correspond to high values of Ed
- c. **low values of the intercept estimate correspond to low values of the slope estimate**
- d. all of the above
- e. None of the above
27. Suppose $\text{wage}_i = \alpha + \beta x_i + \varepsilon_i$ and we have 100 observations on wage and x , with average values 70 and 30, respectively. We have run a regression to estimate the slope of x as 2.0. Consider now a new individual whose age is 20. For this individual the predicted wage from this regression is
- a. **50**
- b. 60
- c. impossible to predict without knowing the intercept estimate
- d. None of the above
28. Suppose you have run the following regression:
- $$y_i = \alpha + \beta x_i + \chi \text{Urban}_i + \delta \text{Immigrant}_i + \phi \text{Urban}_i * \text{Immigrant}_i + \varepsilon_i$$
- where Urban is a dummy indicating that an individual lives in a city ($\text{Urban}=1$) rather than in a rural area ($\text{Urban}=0$), and Immigrant is a dummy indicating that an individual is an immigrant ($\text{Immigrant}=1$) rather than a native. The coefficient δ is interpreted as the ceteris paribus difference in y between:
- a. **an immigrant and a native**
- b. a rural immigrant and a rural native
- c. an urban immigrant and an urban native
- d. an urban immigrant and a rural native
- e. None of the above
29. If the expected value of the error term is 5, then after running an OLS regression

- of y on the relevant explanatory variables and a constant term
- the average of the residuals should be approximately 5
 - the average of the residuals should be exactly zero**
 - the average of the residuals should be exactly five
 - nothing can be said about the average of the residuals
 - None of the above
30. In a logit regression, to report the influence of an explanatory variable x on the probability of observing a one for the dependent variable we report
- the slope coefficient estimate for x
 - the average of the slope coefficient estimates for x of all the observations in the sample
 - the slope coefficient estimate for x for the average observation in the sample
 - none of these**
 - all of these could be a good way to report.
31. To find the maximum likelihood estimates the computer searches over all possible values of the
- dependent variable
 - independent variables
 - coefficients**
 - the independent variables and the coefficients
 - None of the above
32. Consider the multivariate classical linear regression model: $\mathbf{Y} = \mathbf{X}\beta + \epsilon$ where \mathbf{Y} is the $(n \times 1)$ matrix of dependent variables, \mathbf{X} is the $(n \times k)$ matrix of explanatory variables, β is the $(k \times 1)$ matrix of unknown coefficients and ϵ is the $(n \times 1)$ matrix of the random effect on \mathbf{Y} . Which of the following statements is correct?
- Given that \mathbf{C}' is a $(n \times 1)$ vector of constants $[c_1, c_2, \dots, c_n]$, the quantity $\mathbf{C}(\mathbf{Y}\mathbf{Y}')\mathbf{C}'$ is a *quadratic form* in \mathbf{Y} .
 - $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{Y}'\mathbf{X}$ is an unbiased estimator of β .
 - The assumption of *Homoskedasticity alone* implies that the variance-covariance matrix of ϵ is given by $\sigma^2\mathbf{I}_n$ where \mathbf{I}_n is the $(n \times n)$ identity matrix and σ^2 is unknown.
 - The estimator $\hat{\beta}_1$ for the parameter β_1 is a *quadratic form* in the Y values.
 - None of the above**
33. Consider the classical linear regression model: $Y_i = \beta_1 + \beta_2 X_{2i} + \epsilon_i$ In the following set of questions, the parameters β_1 and β_2 are estimated by Ordinary Least Squares (OLS) method to generate estimators $\hat{\beta}_1$ and $\hat{\beta}_2$. Which of the following is a correct statement?
- $\hat{\beta}_2$ will not be an efficient estimator of β_2 if $\text{Cov}(\epsilon_i, \epsilon_j)$ is not zero for some i and j ($i \neq j$).**
 - The covariance of $\hat{\beta}_1$ and $\hat{\beta}_2$ is positive if the mean of the X_{2i} is zero.
 - Suppose that $\text{Cov}(X_{2i}, \epsilon_i) > 0$. The OLS method will produce unbiased estimators provided that the sum of the error terms ϵ_i equals 0.
 - We need to assume that the random term ϵ_i has a normal distribution in order to show that $\hat{\beta}_1$ and $\hat{\beta}_2$ are consistent estimators of β_1 and β_2 respectively.
 - None of the above
34. In the binary dependent variable model, a predicted value of 0.6 means that:
- the most likely value the dependent variable will take on is 60 percent.
 - given the values for the explanatory variables, there is a 60 percent probability that the dependent variable will equal one.**
 - the model makes little sense, since the dependent variable can only be 0 or 1.
 - given the values for the explanatory variables, there is a 40 percent

- probability that the dependent variable will equal one.
- e. None of the above
35. 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
- $\hat{\beta}$ is an unambiguously better estimator than $\tilde{\beta}$.
 - $\tilde{\beta}$ is an unambiguously better estimator than $\hat{\beta}$.
 - $\hat{\beta}$ has to be the OLS estimator.
 - a. an c. are true.
 - None of the above
36. ¿Cuál de las siguientes pruebas para comprobar autocorrelación es válida aun cuando los regresores en la ecuación original incluyen un rezago de la variable dependiente, por ejemplo, en la ecuación: $y_t = \beta_1 + \beta_2 x_t + \beta_3 y_{t-1} + \epsilon_t$?
- Una prueba t de una regresión hecha por MCO entre la variable dependiente y su variable rezaga sin incluir intercepto.
 - La Prueba de Durbin-Watson
 - La Prueba de Breush-Pagan.
 - La prueba de White
 - Ninguna de las anteriores
37. Los estimadores MCO NO serán sesgados si:
- Cualquiera de las variables independientes es medida con error.
 - Una de las variables binarias es medida con error.
 - Se omite una variable independiente que no se encuentra correlacionada con las variables independientes incluidas en la ecuación.
 - Todas las anteriores
 - Ninguna de las anteriores
38. Si el término de error está autocorrelacionado, entonces los coeficientes estimados por el método de MCO son:
- Insesgados y eficientes
 - Insesgados e ineficientes.
 - Sólo ineficientes.
 - Sólo insesgados.
 - Ninguna de las anteriores
39. El problema de simultaneidad implica:
- La posible inexistencia de un sesgo en los estimadores MCO.
 - La necesidad de emplear los estimadores de MCG.
 - La existencia de una variable (explicativa o dependiente) relacionada con el error.
 - Todas las anteriores
 - Ninguna de las anteriores
40. Ante la presencia de un problema de multicolinealidad no perfecta:
- La mejor forma de solucionarla es eliminando una de las variables correlacionadas.
 - En ningún caso es posible solucionarlas.
 - Se considera que los betas están correlacionados con los otros betas y por eso uno solo de ellos recoge todo el efecto de las variables correlacionadas.
 - Todas las anteriores
 - Ninguna de las anteriores.
41. El test de Durbin y Watson es un test de:
- No autocorrelación
 - Autocorrelación
 - Heteroscedasticidad
 - A y B.
 - Ninguna de las anteriores.
42. En un modelo de regresión lineal con una tendencia lineal (time trend), el coeficiente asociado a la tendencia puede ser:
- un estimador de la tasa de crecimiento % de la variable dependiente (ceteris paribus).
 - El cambio en otra de las variables independientes por periodo de tiempo (ceteris paribus).
 - (a) y (b) pueden ser correctas.

- d. Ninguna de las anteriores.
43. Comparando el método de Máxima Verosimilitud (MV) y el de Mínimos Cuadrados Ordinarios para la estimación de un modelo con variable dicotómica como dependiente, NO se puede afirmar que:
- MCO escoge la línea que mejor se ajusta por medio de la minimización de la suma de las desviaciones (en sentido vertical) de cada uno de los valores de Y con respecto a esta línea.
 - MV escoge la línea que mejor se ajusta maximizando la probabilidad conjunta de observar las n-observaciones independientes sobre Y en la muestra.
 - Tendrán resultados diferentes.
 - Todas las anteriores
 - Ninguna de las anteriores.
44. La autocorrelación en los datos es un problema porque:
- El método de MCO asume que los datos no son correlacionados y calcula los estimadores puntuales de los betas de la regresión de manera acorde.
 - Sesga el estimador de la matriz de varianzas y covarianzas.
 - (a) y (b) son correctas.
 - Ninguna de las anteriores
45. Qué supuesto se requiere para que los estimadores de MCO sean insesgados?
- Correr el modelo por MCO es suficiente, ya que el teorema de Gauss-Markov afirma que cualquier especificación de un modelo por MCO dará como resultado estimadores insesgados.
 - Homoscedasticidad: $\text{Var}(\varepsilon_i) = \sigma^2$
 - No autocorrelación: $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ para $i \neq j$.
 - Ninguna de las anteriores
46. La heteroscedasticidad es un problema que puede presentarse en:
- Los modelos que emplean datos de corte transversal.
 - Los modelos que emplean datos de series de tiempo.
 - A y B
- d. Ninguna de las anteriores
47. La diferencia entre las pruebas de autocorrelación de Box-Pierce y Durbin-Watson es:
- La primera detecta solamente autocorrelación de primer orden.
 - La segunda detecta solamente autocorrelación de primer orden.
 - Ambas detectan solamente autocorrelación de primer orden, por lo tanto, éste no sería un factor diferenciador.
 - Ninguna de las anteriores.
48. Los estimadores de máxima verosimilitud para regresiones lineales, dan como resultado:
- Estimaciones insesgadas de los betas.
 - Estimaciones eficientes de la matriz de varianzas y covarianzas.
 - Estimaciones insesgadas tanto de los betas como de la matriz de varianzas y covarianzas.
 - Estimaciones sesgadas tanto de los betas como de la matriz de varianzas y covarianzas.
49. La diferencia entre 8% y 5% es:
- 3 puntos porcentuales
 - 3 por ciento
 - Es indistinto a o b es lo mismo.
 - Dependiendo del contexto en el que se comparen.
 - Ninguna de las anteriores
50. Uno de los siguientes es un supuesto de Gauss-Markov y puede ser relajado y permitir que los estimadores continúen siendo MELI.
- Omisión de variables explicativas no relevantes.
 - X estocásticas.
 - Valor esperado del error igual a 0.
 - X no estocásticas.
 - Ninguna de las anteriores

II. (20 puntos)

Un investigador está interesado en estimar un modelo tipo IS-LM. Usted acaba de ser contratado como asistente de la investigación en reemplazo del asistente que se ira a estudiar al exterior. El anterior asistente estimó una de las ecuaciones simultáneas que conforman el sistema. Estos resultados se reporta al final del examen.

En donde R_t , M_t , Y_t , e Inv_t , corresponden al tipo de interés, la oferta monetaria (en millones de pesos constates de 1994), el PIB (en millones de pesos constates de 1994) y la inversión (en millones de pesos constates de 1994) para el año t, respectivamente. Así mismo, D_t corresponde a una variable dummy que toma el valor de uno para años en los que el Banco Central ha sido independiente de la rama ejecutiva. Para ello dispone de los datos anuales para el periodo 1966-2006 de este país. Teniendo en cuenta esta información, responda:

- a) De acuerdo a los cálculos del econometrista, determine cuáles son las variables endógenas y cuáles las exógenas del sistema. **(5 puntos)**

Variables Endógenas: $\ln(R_t)$ y $\ln(Y_t)$

Variables Exógenas: $\ln(M_t)$, $D_t * \ln(M_t)$, $\ln(R_{t-1})$, $D_t * \ln(R_{t-1})$, $D_t * \ln(Inv_t)$, $\ln(Inv_t)$, w_{t-1} y $D_t * w_t$ (donde $w_t = \ln(Y_t)$)

(0.5 cada una de las variables)

- b) Es posible afirmar que este modelo estimado presenta síntomas de autocorrelación negativa? Explique su respuesta de la manera más clara posible. **(5 puntos)**

De acuerdo a los resultados se puede observar que existe un problema de autocorrelación positiva, dado que el DW es 0.5. Se esperaba que ustedes realizaran la prueba de autocorrelación **(5 puntos)**

- c) Teniendo en cuenta la significancia y omitiendo el problema encontrado (o no), interprete los coeficientes estimados **(5 puntos)**

$\hat{\beta}_1 = 34.99671$ No tiene interpretación económica **(1 punto)**

$\hat{\beta}_2 = 18.7$ La elasticidad del PIB que garantiza el equilibrio en el mercado de bienes y servicios respecto a la inversión. Noten que no es necesario decir que esto es para los años en que el Banco Central no era independiente pues el coeficiente asociado a la dummy no es significativa. **(1 punto)**

$\hat{\beta}_3 = -1.30$ La elasticidad del PIB que garantiza el equilibrio en el mercado de bienes y servicios respecto a la tasa de interés. **(1 punto)**

$\hat{\beta}_4 = -3.096$ NO existe diferencia entre la elasticidad del PIB que garantiza el equilibrio en el mercado de bienes y servicios respecto a la inversión para los años en los que el Banco Central era y no era independiente. **(2 puntos)**

- d) Explique claramente como podría comprobar la hipótesis de que la independencia del Banco de la república ha sido una buena idea para esta economía. **(5 puntos)**

En este caso, se debe emplear un modelo de la forma reducida. Es decir:

$$\ln(Y_t) = \pi_{2,1} + \pi_{2,2} \ln(M_t) + \pi_{2,3} D_t \ln(M_t) + \pi_{2,4} \ln(Inv_t) + \pi_{2,5} D_t \ln(Inv_t) + \pi_{2,6} \ln(R_{t-1}) + \pi_{2,7} D_t * \ln(R_{t-1}) + \pi_{2,8} w_{t-1} + \pi_{2,9} D_t * w_t + \pi_{2,10} D_t + \mu_{2,t}$$

$$\ln(R_t) = \pi_{1,1} + \pi_{1,2} \ln(M_t) + \pi_{1,3} D_t \ln(M_t) + \pi_{1,4} \ln(Inv_t) + \pi_{1,5} D_t \ln(Inv_t) + \pi_{1,6} \ln(R_{t-1}) + \pi_{1,7} D_t * \ln(R_{t-1}) + \pi_{1,8} w_{t-1} + \pi_{1,9} D_t * w_t + \pi_{1,10} D_t + \mu_{1,t}$$

(2 puntos)

Por otro lado, es importante tener en cuenta que es necesario determinar inicialmente si existe un cambio estructural o no. En este caso esto implica probar para cada una de las ecuaciones de la forma reducida las siguientes restricciones:

$$\pi_{2,3} = \pi_{2,5} = \pi_{2,7} = \pi_{2,9} = \pi_{2,10} = 0$$

$$\pi_{1,3} = \pi_{1,5} = \pi_{1,7} = \pi_{1,9} = \pi_{1,10} = 0$$

Esto se corresponde a restricciones de la forma $R\beta = c$. (ya hemos discutido ampliamente como comprobar esto) (Se esperaba que los estudiantes plantearan estas dos hipótesis y mostraran como comprobarlas (2 puntos por esto, uno por cada una):

En general, no es muy claro cuál debería ser el signo de asociadas a las variables Dummy. Sin embargo si es deseable que $\pi_{2,9} < 0$.

III. (30 puntos)

Un investigador está interesado en estimar un modelo que permita determinar las variables que afectan el hecho de que un individuo sea jugador de chance (J_i es una variable dummy que toma el valor de uno si el individuo es jugador de chance y cero en caso contrario) y el monto de dinero que los individuos apuestan en cualquier juego de azar en el Valle del Cauca (A_i representa el gasto mensual en juegos de azar del individuo i en pesos para el juego de chance).

Para lograr su objetivo, el investigador emplea una base de datos que corresponde a encuestas realizadas a una muestra representativa de jugadores en el Valle del Cauca. Así mismo, emplea las dos siguientes especificaciones:

$$\ln(A_i) = \alpha_1 + \alpha_2 \ln(Y_i) + \alpha_3 G_i + \alpha_4 \ln(edu_i) + \alpha_5 J_i + \varepsilon_i \tag{1}$$

$$J_i = \beta_1 + \beta_2 G_i + \beta_3 \ln(edu_i) + \mu_i \tag{2}$$

donde Y_i , G_i , y edu_i corresponden a al ingreso del individuo i (medio en miles de pesos por mes), una variable dummy que toma el valor de uno si el individuo i es un hombre y cero en caso contrario y los años de educación del individuo i , respectivamente.

a) Rápidamente, discuta que tipo de problema(s) econométrico(s) puede esperar (a priori) el investigador tener si estima (1) y (2) por MCO. **Sea lo más concreto y preciso posible (10 puntos)**

Noten que estas dos ecuaciones se consideran ecuaciones simultaneas, donde J_i y $\ln(A_i)$ son las variables endógenas.

- La primera ecuación presentará el problema de Simultaneidad (**2 puntos**) y de pronto problemas de heteroscedasticidad dado el tipo de datos que se emplean (v.g. corte transversal) (**1 punto**). Sin embargo, el segundo problema es trivial dado que esta ecuación está subidentificada y por tanto no es posible estimarla (**2 puntos**).
- La segunda tendrá problema de heteroscedasticidad, dado que esto corresponde al modelo de probabilidad lineal (**2 puntos**). En especial se espera que $\sigma_i^2 = J_i(1 - J_i)$. (**2 puntos**) Y no hay problema de simultaneidad pues no hay variable endógena como explicativa en este modelo. (**1 punto**)

Ignorando los problemas (o no) que se pueden presentar en la especificación (2), se recoge la información necesaria para estimarla por medio del método de MCO. Para esto el asistente del investigador preparó las siguientes matrices:

$$X^T X = \begin{bmatrix} 2500 & 500 & 0 \\ 500 & 200 & 0 \\ 0 & 0 & 12 \end{bmatrix} \quad X^T y = \begin{bmatrix} 300 \\ 200 \\ 180 \end{bmatrix}$$

- b) De acuerdo a estos datos, determine (si es posible) cuál es el tamaño de la muestra empleada, cuál es la proporción de NO jugadores de chance según la muestra y que proporción de la muestra corresponde a mujeres. (**5 puntos**)

De las dos anteriores matrices se puede determinar que:

- el tamaño de la muestra empleada = $n = 2500$ (**1 punto**)
- la proporción de jugadores según la muestra = $3/25 = 12\%$ Por tanto la proporción de no jugadores de chance es 88% (**2 puntos**)
- proporción de la muestra corresponde a mujeres = $1 - (500/250) = 80\%$ (**2 puntos**)

- c) Encuentre los estimadores MCO del modelo (2). (**10 Puntos**)

En este caso tenemos que:

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

$$X^T X = \begin{bmatrix} 0 & -1/500 & 0 \\ -1/500 & 1/100 & 0 \\ 0 & 0 & 1/12 \end{bmatrix} \begin{bmatrix} 300 \\ 200 \\ 180 \end{bmatrix} = \begin{bmatrix} -4/25 \\ 7/5 \\ 15 \end{bmatrix} = \begin{bmatrix} -0.2 \\ 1.4 \\ 15 \end{bmatrix}$$

- d) Interprete el significado de cada uno de los coeficientes estimados. (**5 Puntos**)

$\hat{\beta}_1 = -0.2$. Una interpretación al pie de la letra sería: La probabilidad de jugar chance para las mujeres que no tiene un año de educación es de -0.2. (A la interpretación anterior no se le asigne ningún puntaje)

Pero no tiene sentido tener una probabilidad negativa y es mejor la siguiente interpretación: La probabilidad de jugar chance para las mujeres con un año de educación es de cero. (**1 punto**)

$\hat{\beta}_2 = 1.4$ otra vez una interpretación ingenua sería: Los hombres tienen en promedio una probabilidad de jugar chance 140 puntos porcentuales mayor que las mujeres. (A la interpretación anterior no se le asigne ningún puntaje)

Esto no tiene sentido y por tanto sería mejor decir que: los hombres juegan con seguridad (2 *puntos*)

$\hat{\beta}_3 = 15$ Un aumento del 1% en los años de educación aumenta la probabilidad de jugar chance en 15 (15*100/100) puntos porcentuales. (2 *puntos*)

Resultados de EasyReg.

Two-stage least squares:

Dependent variable:

$$Y = \ln(Y)$$

X variables, including instrumental variables:

$$X(1) = \ln(INV)$$

$$X(2) = \ln(M)$$

$$X(3) = \ln(R)$$

$$X(4) = \text{LAG1}[\ln(Y)]$$

$$X(5) = \ln(\text{LAG1}[R])$$

$$X(6) = D * \ln(INV)$$

$$X(7) = D * \ln(M)$$

$$X(8) = D * \ln(\text{LAG1}[R])$$

$$X(9) = D * (\text{LAG1}[\ln(Y)])$$

$$X(10) = 1$$

Endogenous X variable:

$$Y^*(1) = \ln(R)$$

Exogenous X variables:

$$X^*(1) = \ln(INV)$$

$$X^*(2) = D * \ln(INV)$$

$$X^*(2) = 1$$

2SLS estimation results for $Y = Y$

Variables	2SLS estimate	t-value	[p-value]
$\ln(INV)$	18.6928	5.339	[0,00000]
$\ln(R)$	-1.30960	-3.903	[0,00009]
$D * \ln(INV)$	3.0960	4.903	[0,1009]
1	34.99671	6.118	[0,00000]

[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-square = 0.80

Effective sample size (n) = 40

DW = 0.5

