ECONOMETRICS - FINAL EXAM, 3rd YEAR (GECO & GADE)

June 28, 2017 – 15:30

First family name:	**Second family Name:**
Name:	**GECO/GADE:**
Mobile:	**E-mail:**

Question 1	**A**	**B**	**C**	**Blank**
Question 2	**A**	**B**	**C**	**Blank**
Question 3	**A**	**B**	**C**	**Blank**
Question 4	**A**	**B**	**C**	**Blank**
Question 5	**A**	**B**	**C**	**Blank**
Question 6	**A**	**B**	**C**	**Blank**
Question 7	**A**	**B**	**C**	**Blank**
Question 8	**A**	**B**	**C**	**Blank**
Question 9	**A**	**B**	**C**	**Blank**
Question 10	**A**	**B**	**C**	**Blank**
Question 11	**A**	**B**	**C**	**Blank**
Question 12	**A**	**B**	**C**	**Blank**
Question 13	**A**	**B**	**C**	**Blank**
Question 14	**A**	**B**	**C**	**Blank**
Question 15	**A**	**B**	**C**	**Blank**
Question 16	**A**	**B**	**C**	**Blank**
Question 17	**A**	**B**	**C**	**Blank**
Question 18	**A**	**B**	**C**	**Blank**
Question 19	**A**	**B**	**C**	**Blank**
Question 20	**A**	**B**	**C**	**Blank**

Correct	**Incorrect**	**Blank**	**Final grade**

INSTRUCTIONS

This exam includes 20 multiple choice questions.

Your answers must be marked on the answer sheet that you will find in the first page. If you want to leave any question unanswered, choose the "Blank" option. This answer sheet is the only part of this exam that will be graded.

A correct answer adds 2 points to the final grade while an incorrect one subtracts 1 point. A blank answer does not add or subtract. The final grade is the number of points divided by 4.

Make sure that you checked your options, including “Blank”. Do not unclip the sheets. Use the blank space in the following pages to write notes or to do arithmetic calculations.

YOU HAVE ONE HOUR AND A HALF TO ANSWER THIS TEST

REMINDER

YOU ARE NOT ALLOWED TO USE DEVICES WITH CONNECTIVITY TO THE INTERNET, INCLUDING MOBILE PHONES, TABLETS, SMARTWATCHES OR MP3/4 PLAYERS

Question 1. Consider the following regression models, where "ln" denotes the natural log:

(1) $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i$

(2) $\ln y_i = \beta_2 \ln x_{i2} + \beta_3 \ln x_{i3} + v_i$

(3) $\ln y_i = \beta_1 + \beta_2 \ln x_{i2} + \beta_3 \ln x_{i3} + w_i$

Which of the following statements is **TRUE**?

A) Under the standard hypotheses of the linear regression model, the sum of OLS residuals in models (1), (2) and (3) is zero

B) The R-squared corresponding to the OLS estimation of model (1) is always larger than that of model (3)

C) If $x_{i2} = 4$ for all i in the three models above, there will be exact collinearity in models (1) and (3)

Question 2. Consider the following regression models, estimated by OLS: (1) $y_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{u}_i$ and (2) $x_i = \hat{\gamma}_1 + \hat{\gamma}_2 y_i + v_i$, where the lineal correlation coefficient between y_i and x_i is 0.75 and the sample standard deviations of both variables are $\hat{\sigma}_y = \hat{\sigma}_x = 1$. Under these conditions, which of the following statements is **FALSE**?

A) The R-squared in both models is the same.

B) The OLS estimates $\hat{\beta}_2$ and $\hat{\gamma}_2$ are such that $\hat{\beta}_2 = \frac{1}{\hat{\gamma}_2}$

C) The OLS estimates $\hat{\beta}_2$ and $\hat{\gamma}_2$ are such that $\hat{\beta}_2 = \hat{\gamma}_2 = 0.75$

Question 3: Under the standard hypotheses of the General Linear Model $Y = X\beta + \varepsilon$, efficiency in the Gauss-Markov sense of the OLS estimator of β means that:

A) The variance-covariance matrix of the OLS estimator of β is the identity matrix $\boldsymbol{I}$

B) There is no alternative linear and unbiased estimator of β with a smaller variance

C) The expected value of the OLS estimator of β is always equal to zero.

Question 4: In the model $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, with $i = 1,2,\ldots,n$, the errors are heteroscedastic when:

A) $var(\varepsilon_i) = \sigma_i^2$ for all $i = 1,2,\ldots,n$

B) $\varepsilon_i = 10 + u_i$ with $var(u_i) = 5$ for all $i = 1,2,\ldots,n$

C) $var(\varepsilon_i) = 2\sigma^2$ for all $i = 1,2,\ldots,n$

Question 5: Consider the model $E_i = \beta_1 + \beta_2 D_i + \varepsilon_i$, where E_i denotes the number of years of schooling for the *i*-th individual and D_i is a dummy variable which value is 1, if the individual lives in an urban zone, and 0 otherwise. The average number of years of schooling is 15 and 10, for individuals living in an urban zone (150 cases in the sample) and in a non-urban zone (150 cases), respectively. In these circumstances, the OLS estimates for the parameters β_1 and β_2 in the previous model are:

A) 15 and 10, respectively

B) 10 and 5, respectively

C) 15 and 5, respectively

Questions 6 to 8 refer to the following statement. **Table 1** summarizes the main estimation results for the model:

$$C_t = \beta_1 + \beta_2 R_t + \beta_3 \mathrm{I}_t + u_t$$

…where (*C*) denotes total household Consumption in Spain, (*R*) is the national Income, and (*I*) denotes the Taxes collected from the private sector. The time span goes from 1974, 1st quarter, to 1998, 4th quarter and all the variables are expressed in million euros.

Table 1

Dependent variable: C_t **Sample:** I/1974 to IV/1998 (T=100)			
	Coefficient	Std. Error	t-Statistic
Constant	145	40	3.63
R_t	0.84	0.26	3.23
I_t	-0.53	0.22	-2.41

R-squared	0.76	$\text{cov}(\hat{\beta}_2 \hat{\beta}_3) = -0.05$ $\text{Prob}[F(1,97) \leq 3.939] = 0.95$
Residual std. dev.	1.425	$\text{Prob}[t(97) \leq -1.661] = 0.05$

Question 6. Given the information in Table 1 and the null that the marginal propensity to consumption (β_2) is one, against the alternative that it is less than one:

A) There is not enough information to test this hypothesis.

B) The test statistic can be computed and the null is rejected in favor of the alternative hypothesis with a 5% significance.

C) The test statistic can be computed and the null is not rejected in favor of the alternative hypothesis with a 5% significance.

Question 7. The test for the null that disposable income (that is: Income – Taxes) is the single explanatory variable for Consumption, instead of Income and Taxes separately, can be written as:

A) $H_{0'}: \beta_3 = -\beta_2$

B) $H_{0'}: \beta_3 = \beta_2$

C) $H_{0'}: \beta_3 + \beta_2 = -1$

Question 8: Which of the following statements is **TRUE**:

A) The test stated in the previous question can be implemented by estimating by OLS both, the model in Table 1 and the model $C_t = \beta_1 + \beta_2(\text{R}_t - \text{I}_t) + v_t$. If the R-squared in this model were 0.67, the null would be rejected at the 5% significance level.

B) The test stated in previous question can be implemented by estimating by OLS both, the model in Table 1 and the model $C_t = \beta_1 + \beta_2(\text{R}_t - \text{I}_t) + v_t$. If the R-squared in this model were 0.67, the null would not be rejected at the 5% significance level.

C) The test stated in previous question can be implemented by estimating by OLS both, the model in Table 1 and the model $C_t = \beta_1 + \beta_2(\text{R}_t + \text{I}_t) + v_t$. If the R-squared in this model were 0.67, the null would not be rejected at the 5% significance level.

Question 9. Consider the model $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, which satisfies the standard hypotheses, but $\text{var}(\varepsilon_i) = 1/z_i^2$ for all i=1,2,…,N. Under these conditions, an efficient estimate of the β_2 parameter could be obtained by estimating by OLS the model:

A) $y_i\, z_i = \beta_1\, z_i + \beta_2\, x_i\, z_i + v_i$ with $\text{var}(v_i) = \text{var}(\varepsilon_i) = 1/z_i^2$ for all i=1,2,…, N

B) $y_i\, z_i = \beta_1 + \beta_2 x_i\, z_i + v_i$ with $\text{var}(v_i) = 1$ for all i=1,2,…, N

C) $y_i\, z_i = \beta_1\, z_i + \beta_2 x_i\, z_i + v_i$ with $\text{var}(v_i) = 1$ for all i=1,2,…, N

Questions 10 to 17 refer to the following statement. We fitted a model to explain the (log) markup (l_MARKUP) of 554 bank branches as a function of: 1) the branch (log) turnover[1] (l_TURNOVER); 2) the (log) number of clients of the branch (l_CLIENTS); 3) the salesforce, measured as the number of employees in the branch (SALES_FORCE) and 4) its SQUARE (SQ_SALES_FORCE), as well as five dummy variables, denoted by Cj (j=1, 2, 3, 4 y 5), so the value of Cj is one if the i-th branch has the complexity level j, and zero otherwise. Bear in mind that C1 and C5 correspond, respectively, to the highest and smallest degrees of complexity.

Model 1
Dependent variable: l_MARKUP
OLS, using observations 1-554

	Coefficient	Standard error	t-statistic	p-value
Constant	2.1149	0.3302	6.4043	<0.0001
l_TURNOVER	0.8198	0.0392	20.8942	<0.0001
l_CLIENTS	0.3144	0.0409	7.6795	<0.0001
SALES_FORCE	0.0161	0.0333	0.4829	0.6294
SQ_SALES_FORCE	−0.0024	0.0023	−1.0330	0.3021
C2	0.1145	0.0352	3.2482	0.0012
C3	0.0831	0.0442	1.8816	0.0604
C4	0.0929	0.0530	1.7533	0.0801
C5	0.0812	0.0636	1.2772	0.2021

[1] In Spanish: "volumen de negocio"

Mean dependent var.	11.7720	S.D. dependent var.	0.7946
Sum squared resid.	35.4814	S.E. of regression	0.2552
R-squared	0.8984	Adjusted R-squared	0.8969
F(8, 545)	602.23	P-value(F)	0.0000

Question 10. The errors in **Model 1** could be heteroscedastic, so we run an auxiliary regression of the squared residuals from Model 1 on: (a) a constant term, (b) the regressors in Model 1, (c) their squared values and (d) their cross-products. If the R^2 of this regression is 0.393 and $\text{Prob}\left(\chi^2_{33} > 217.685\right) = 0$, then:

A) …the null of homoscedasticity would be rejected at any significance level.

B) …the null of homoscedasticity would not be rejected at the 10% significance level.

C) …there is not enough information in the problem statement to test the null of homoscedasticity.

Question 11. The heteroscedasticity test used in previous question is known as:

A) …Jarque-Bera test.

B) …Breusch-Pagan test.

C) …White test.

Question 12. The value of the Jarque–Bera statistic corresponding to the OLS residuals of **Model 1** is 161.134. If $\text{Prob}\left(\chi^2_2 > 5.99\right) = 0.05$, then:

A) …the null of normality should be rejected at the 5% significance level.

B) …the null of normality cannot be rejected at the 5% significance level.

C) …the Jarque-Bera statistic in Model 1 follows a χ^2_3 distribution under the null.

According to the results of the previous questions we re-estimated Model 1 using White's heteroscedasticity-robust standard errors for the parameters. Model 2 summarizes some results of this exercise.

Model 2:

Dependent variable: l_MARKUP

Using observations 1-554

Heteroscedasticity-robust standard errors

	Coefficient	Standard error	t-statistic	p-value
Constant	2.1149	0.5026	4.2076	<0.0001
l_TURNOVER	0.8198	0.0392	20.9040	<0.0001
l_CLIENTS	0.3144	0.0727	4.3261	<0.0001
SALES_FORCE	0.0161	0.0312	0.5154	0.6063
SQ_SALES_FORCE	−0.0024	0.0020	−1.1652	0.2440
C2	0.1145	0.0333	3.4323	0.0006
C3	0.0831	0.0442	1.8780	0.0604
C4	0.0929	0.0568	1.6357	0.1019
C5	0.0812	0.0575	1.4128	0.1577

Question 13. According to the information in **Models 1 and 2**, the 95% confidence interval for the coefficient associated to l_CLIENTS is (use all the available decimals in your calculations and assume that the value of a Student t with 545 degrees of freedom which leaves in both tails a 5% probability is 2):

A) [0.3144 , 0.4598]

B) [0.1690 , 0.4598]

C) [0.1690 , 0.3871]

Question 14. According to the results in **Model 2**:

A) The differential effect on MARKUP of a level C5 branch, in comparison with a C1 branch, assuming the same values for the rest of the variables, is approximately 8.12% and is significant at 10% level.

B) The differential effect on MARKUP of a level C2 branch, in comparison with a C1 branch, assuming the same values for the rest of the variables, is approximately 11.45% % and is significant at 5% level.

C) The differential effect on MARKUP of a level C4 branch, in comparison with a C1 branch, assuming the same values for the rest of the variables, is approximately 9.29% % and is significant at 5% level.

Question 15: According to the results in **Model 2**, if we want to test the null that the Coefficient associated to a level C3 branch is equal to that of a C4 branch, and knowing that $\text{Prob}[t(545) \leq -2] = 0.025$:

A) The null must be rejected at the 5% significance level.

B) The null cannot be rejected at the 5% significance level.

C) There is not enough information in the problem statement to test this null.

Question 16: According to the results in **Models 1 and 2**, which of the following statements are **FALSE:**

1. Regardless of the results of the residual heteroscedasticity and normality tests, the model with a larger R-squared or smaller Akaike Information Criterion should always be preferred.
2. Any hypotheses testing or interval forecast for MARKUP must be done using Model 2 results.
3. If the errors in Model 1 are heteroscedastic, the OLS estimator remains efficient in the Gauss-Markov sense, but is no longer unbiased.

A) Affirmations 1 and 2 are false.

B) Affirmations 2 and 3 are false.

C) Affirmations 1 and 3 are false.

Question 17: According to the results in **Model 2**, the l_MARKUP forecast of a C3 branch, with SALES_FORCE = 10 and l_TURNOVER = l_CLIENTS = 100 is:

A) 113.424

B) 115.539

C) 115.779

Question 18: Consider a monthly time series (Y_t) from January 1998 to December 2010. If we denote $\nabla Y_t = Y_t - Y_{t-1}$, $\nabla_{12} Y_t = Y_t - Y_{t-12}$ and "ln" is the natural (neperian) logarithm:

A) The series $\nabla \ln Y_t$ can be interpreted as an absolute growth indicator for Y_t.

B) The series $\nabla\nabla_{12} \ln Y_t$ can be interpreted as the logarithmic growth rate of Y_t.

C) The series $\nabla_{12} \ln Y_t$ can be interpreted as the logarithmic growth rate of Y_t accumulated during the last 12 months or, equivalently, the annual logarithmic growth rate.

Question 19. Figure 1 displays the number of cases of Melanoma (M_t) in the male population of a US State, from 1936 to 1972. Figure 2 shows the Spanish Industrial Production Index (IPI_t), from January 1975 to March 2001.

Figure 1: M_t

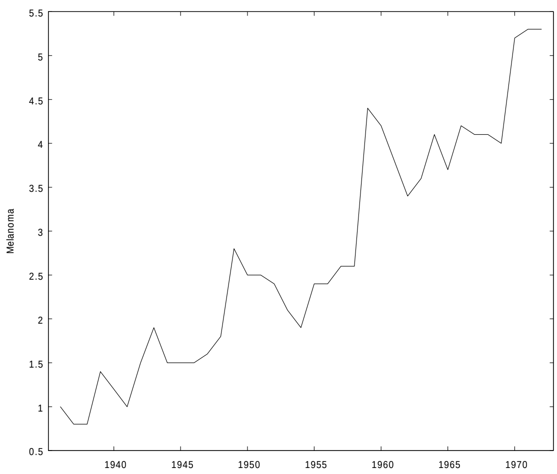
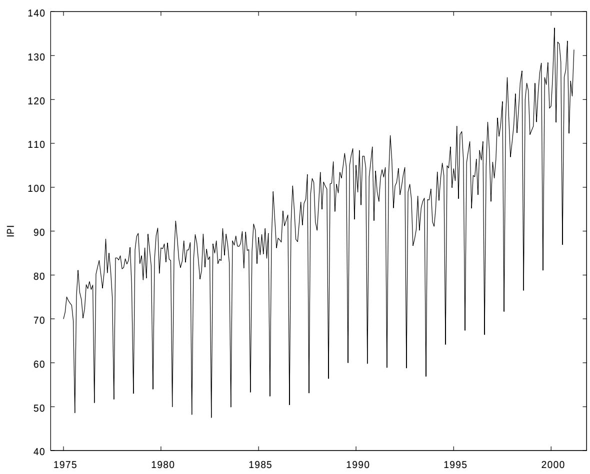


Figure 2: IPI_t



According to figures 1 and 2, which of the following statements is **FALSE**:

A) IPI_t displays a strong seasonality.

B) Both time series are mean stationary.

C) Both time series display a trending behavior.

Question 20. Consider a linear regression model relating time series data. Which of the following instruments would **NOT** be suitable to detect autocorrelation in the error term:

A) White test.

B) A time series plot of the residuals.

C) Estimating by OLS a new model differencing both, the endogenous and explanatory variables, and then testing the residuals for mean stationarity.

Calculations

ECONOMETRICS - FINAL EXAM, 3rd YEAR (GECO & GADE)

June 28, 2017 – 15:30

First family name:	**Second family Name:**
Name:	**GECO/GADE:**
Mobile:	**E-mail:**

Question 1	**A**	**B**	**C**	**Blank**
Question 2	**A**	**B**	**C**	**Blank**
Question 3	**A**	**B**	**C**	**Blank**
Question 4	**A**	**B**	**C**	**Blank**
Question 5	**A**	**B**	**C**	**Blank**
Question 6	**A**	**B**	**C**	**Blank**
Question 7	**A**	**B**	**C**	**Blank**
Question 8	**A**	**B**	**C**	**Blank**
Question 9	**A**	**B**	**C**	**Blank**
Question 10	**A**	**B**	**C**	**Blank**
Question 11	**A**	**B**	**C**	**Blank**
Question 12	**A**	**B**	**C**	**Blank**
Question 13	**A**	**B**	**C**	**Blank**
Question 14	**A**	**B**	**C**	**Blank**
Question 15	**A**	**B**	**C**	**Blank**
Question 16	**A**	**B**	**C**	**Blank**
Question 17	**A**	**B**	**C**	**Blank**
Question 18	**A**	**B**	**C**	**Blank**
Question 19	**A**	**B**	**C**	**Blank**
Question 20	**A**	**B**	**C**	**Blank**

Correct	**Incorrect**	**Blank**	**Final grade**