ISEG, Univeridade de Lisboa
Econometrics EXAM
January 31, 2023

TOTAL TIME 2 HOURS

Instructions

The exam has 2 parts: **Section A** and **B**, each with a **50% weight** in the final grade. Your final grade has to be 9.5 to pass this exam (irrespective of the mark in each section and irrespective of the mark from January 5, 2023 exam). During this exam, you are allowed to **only** use the formulae sheet provided and a calculator (but not tablets or phones). Lecture notes or books are **NOT** allowed.

SECTION A

Total marks 20 (50% weight in the final grade).

Solve both problems 1 and 2 below.

Problem 1. Total 3 marks

Answer the following three multiple choice questions. Do not give a justification of your answer. There is only one correct answer for each question: (a), (b) or (c). Each correct answer receives a mark of 1. Each incorrect answer receives a penalty of -0.25.

I. One of the main assumptions in econometrics is $E(u|x) = 0$, where u is the error term and x is the regressor. This assumption implies:

(a) $E(x|u) = 0$.
(b) $E(u) = 0$.
(c) $E(u^2) = 0$

II. Consider the following linear regression model: $y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 z_t + u_t$. Let $x_t = 2z_t + v_t$, where v_t is i.i.d. with mean zero and variance 1. Denote by $\hat{\alpha}_i$, $i = 1, 2, 3$, the Ordinary Least Squares (OLS) estimators.

(a) x_t and z_t are perfectly colinear, so one of the assumptions to prove unbiasedness of the OLS estimators is violated.

(b) x_t and z_t are not perfectly colinear, so one of the assumptions to prove unbiasedness of the OLS estimators is not violated.

(c) The OLS estimators are unbiased whether or not x_t and z_t are perfectly multicolinear.

III. Consider the following linear regression model: $y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 z_t + u_t$. The estimation results from this linear regression model can be used to compute a t-statistic to test one of the following null hypotheses:

(a) $\alpha_0 = \alpha_1 = \alpha_2$

(b) $\alpha_0 = \alpha_1 = \alpha_2 = 0$

(c) $\alpha_2 = 0$

Problem 2. Total 17 marks

Let *return* be the total return from holding a firm's stock over the four-year period from the end of 2018 to the end of 2022. The efficient markets hypothesis says that these returns should not be systematically related to information known in 2018. If firm characteristics known at the beginning of the period help to predict stock returns, then we could use this information in choosing stocks. For 2018, let *dkr* be a firm's debt to capital ratio, let *eps* denote the earnings per share, let *lnetinc* denote the logarithm of net income, and let *salary* denote total compensation for the CEO. The following equation is considered to verify the efficient market hypothesis:

$$return = \delta_0 + \delta_1\ dkr + \delta_2\ eps + \delta_3\ lnetinc + \delta_4\ salary\ + u \qquad (1)$$

with the estimation results given in the following Stata output:

Source	SS	df	MS			
				Number of obs	=	142
				F(4, 137)	=	1.54
Model	9437.52612	4	2359.38153	Prob > F	=	0.1936
Residual	209658.651	137	1530.35512	R-squared	=	0.0431
				Adj R-squared	=	0.0151
Total	219096.178	141	1553.8736	Root MSE	=	39.12

return	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
dkr	.3374942	.2017106	1.67	0.097	-.0613745	.736363
eps	.06067	.0801067	0.76	0.450	-.0977356	.2190755
lnetinc	-4.017061	3.068515	-1.31	0.193	-10.08484	2.050717
salary	.0036601	.0022008	1.66	0.099	-.0006917	.008012
_cons	4.885508	17.23437	0.28	0.777	-29.19426	38.96528

(a) Interpret the estimate of the δ_3 coefficient.

(1 Mark)

(b) The column entitled "t" (from t-statistic) is missing. Use the relevant information in the Stata output and compute this column.

(2 Marks)

(c) Write the relevant null and alternative hypotheses for which the missing t-statistics are used.

(2 marks)

(d) Using your answer from (b) and the relevant critical values for a 10% significance level, decide whether or not you reject the null hypotheses from (c). Justify your answer. Based on your answer, can you say whether *dkr*, *eps*, *lnetinc* and *salary* are individually significant?

(2 marks)

(e) Do you reach the same conclusion as in (c) if you use the p-values from column four of the Stata output? Justify your answer using a 10 % significance level.

(2 Marks)

(f) In the top right panel of the Stata output, the value of the F-statistic is missing, as well as the degrees of freedom $F(\cdot, \cdot)$. Fill in the missing values.

(2 Marks)

(g) State the null and alternative hypothesis for which the F-statistic in the Stata output is meant for. Give an interpretation of it.

(1 Mark)

(h) You are interested to see if *roe* (return on equity) and *rok* (return on capital) have no effect on *return* (after controlling for *dkr*, *eps*, *lnetinc*, *salary*). Write the restricted regression and the unrestricted regression. The sum of squared residuals (SSR) from one of the regressions is 200000. Using this SSR and the Stata output above compute the relevant statistic that would allow you to decide whether *roe* and *rok* have no effect on *return*.

(4 Marks)

(i) Give an interpretation of the 95% confidence interval corresponding to the parameter associated with *eps* in the Stata output above.

(1 mark)

SECTION B

Total 20 marks (50% weight in the final grade)

Solve Problems 3, 4 and 5 below.

Problem 3. Total 3 marks

Answer the following multiple choice questions. Do not give a justification of your answer. There is only one correct answer. Each correct answer receives a mark of 1. Each incorrect answer receives a penalty of -0.25.

I. A financial analyst considers the following linear regression model: $y_t = \alpha_0 + \alpha_1 x_t + u_t$ to explain the returns of a stock. However, the true model is $y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 z_t + u_t$. We also know that the correlation between x_t and z_t is 0.7 and $\alpha_2 = 0.5$.

(a) Then the OLS estimator $\tilde{\alpha}_1$ from the model considered by the analyst is known to be biased and the direction of the bias is negative.

(b) Then the OLS estimator $\tilde{\alpha}_1$ from the model considered by the analyst is known to be biased and the direction of the bias is positive.

(c) Then the OLS estimator $\tilde{\alpha}_1$ from the model considered by the analyst is not biased despite the missing regressor z_t.

II. Consider the time series process: $y_t = \rho y_t + u_{t-1}$, $t = 1, 2, \ldots$, where v_t is a white noise.

(a) y_t is a MA(1) and is covariance-stationary only if $|\rho| < 1$.

(b) y_t is an AR(1) process and is covariance-stationary only if $|\rho| < 1$

(c) y_t is a finite distributed lag model of order 1.

III. Consider the following linear regression model: $y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 z_t + u_t$, $t = 1, \ldots, n$, and n is the sample size. Denote by $\hat{\alpha}_i$, $i = 1, 2, 3$, the Ordinary Least Squares (OLS) estimators. Suppose that for the data at hand, in this model the assumption that $Var(u_t)$ =constant for all t is not satisfied. The fact that $Var(u_t) \neq$constant means that:

(a) The errors u_t, $t = 1, \ldots, n$ are autocorrelated, and as a consequence the OLS estimators are biased.

(b) The errors u_t, $t = 1, \ldots, n$ are homoskedastic and as a consequence the OLS estimators are biased.

(c) The errors u_t, $t = 1, \ldots, n$ are heteroskedastic, but this by itself does not imply that the OLS estimators are biased.

Problem 4. Total 11 marks

Consider the model:

$$\begin{aligned}\log(wage) &= \gamma_0 + \gamma_1\, male + \gamma_2\, educ + \gamma_3\, male \cdot educ \\ &+ \gamma_4 exper + \gamma_5 exper^5 + u\end{aligned}$$

where *educ* is education (in years), *male* is a dummy variable that takes the value 1 for men and zero otherwise, amd *exper* is experience (in years). The estimated model is:

$$\begin{aligned}\log(\widehat{wage}) &= \underset{(0.119)}{0.389} + \underset{(0.168)}{0.227}\, male + \underset{(0.008)}{0.082}\, educ + \underset{(0.0131)}{0.056}\, male \cdot educ \\ &+ \underset{(0.005)}{0.029} exper - \underset{(0.00011)}{0.00058} exper^2\end{aligned}$$

with the standard errors in between the round brackets.

(a) What is the estimated return to education for men and for women? (1 mark)

(b) Is the return to education the same for men and for women? Give a detailed answer. To answer this question use a 5% significance level.

(2 marks)

(c) Interpret the parameter estimate associated with male.

(1 mark)

(d) What other regression do you need to run to test the null hypothesis that, holding the other factors fixed, exper has no effect on wage?

(1 mark)

(e) Consider the Lagrange Multipler (LM) test for testing the restriction in (d). What are the steps to construct the LM test statistic?

(3 marks)

(f) Find the value of *educ* such that the predicted values of $\log(wage)$ are the same for men and women.

(3 marks)

Problem 5. Total 6 marks

An analyst from the Bank of Portugal is considering the following two models for explaining the Gross Domestic Product (GDP), denoted here by z_t:

$$z_t = \alpha_0 + \alpha_1 t + v_t \quad (2)$$

$$z_t = \rho z_{t-1} + u_t \quad (3)$$

where v_t and u_t are i.i.d.(0,1), and $\rho = 1$.

(a) What is the name of each model?

(1 mark)

(b) Derive the unconditional mean and variance of z_t implied by each model above.

(4 marks)

(c) Is any of the two models covariance stationary or/and weak dependent? Briefly justify your answer.

(1 mark)

END OF EXAM