**UNIVERSITY OF DUBLIN**

**TRINITY COLLEGE**

**FACULTY OF ARTS, HUMANITIES AND SOCIAL SCIENCES**

**DEPARTMENT OF ECONOMICS**

**Junior Sophister Michaelmas Term 2013**

**Supplemental Examination**

**EC3090 Econometrics**

**DATE VENUE TIME**

Professors O. Henry, C. Newman, A. Bénétrix and Mr M. Curran

EXAM INSTRUCTIONS

Section A: Please answer Q1 **and** one of **either** Q2 **or** Q3

Section B: Please answer TWO from Q4, Q5, Q6, Q7

**Materials Permitted for this Examination**

Standard non-programmable calculator

**You may not start this examination until you are instructed to do so by the Invigilator.**

**Section A**

**Please answer Q1 and EITHER Q2 OR Q3**

**Question 1 (100 marks)**

a) Define what is meant by the term heteroscedasticity in the context of the general linear regression model using a real life application of your choice to illustrate. What are the consequences for Ordinary Least Squares Estimation of the general linear model when heteroscedasticity is present?

*[25 marks]*

b) Let  be a random sample from a population with a normal distribution with mean  and variance . Consider the following two alternative estimators of :

 and 

Compare the small-sample properties of these estimators.

*[25 marks]*

c) Consider the following linear regression model:



Outline how you would construct an F-test of the null hypothesis that  clearly stating any assumptions required.

*[25 marks]*

d) Consider the following estimated wage regression equation (standard errors in parenthesis) for a sample of 1,000 adults:





Where,  is the wage level earned by individual *i* and  is the years of schooling attained by individual *i*. Interpret the coefficient on  and comment on its statistical significance. Are you confident that estimating this model using Ordinary Least Squares produces an unbiased estimate of the relationship between years of schooling and wages? Explain your answer.

*[25 marks]*

**Question 2 (100 marks)**

Consider the following population model:



a) Explain what is meant by Ordinary Least Squares (OLS) estimation and outline how you would use this approach to estimate the parameters of this model. *[20 marks]*

b) What assumptions are required for the parameters of this model to be unbiased? Explain in a practical sense what each of these assumptions means using examples. *[20 marks]*

c) Derive a test of the null hypothesis that  against the two sided alternative clearly stating any distributional assumptions required in the construction of the test statistic. *[20 marks]*

d) What are the consequences for OLS estimation if  and . Explain what this means in practice using an example. *[20 marks]*

e) What are the consequences for OLS estimation if  and . Explain how you would proceed with the estimation of this model in practice. *[20 marks]*

**Question 3 (100 marks)**

Consider the following equation to explain the total amount a household has borrowed in the form of loans from banks () (expressed in euros) in terms of income ():



a) What sign might you expect on the parameter ? Explain your answer. *[10 marks]*

b) Give examples of two other control variables that you think are important to include in this model and explain why. What sign might you expect on the parameters associated with these variables? Explain your answer. *[20 marks]*

c) Using data on 1,500 households the following equation was estimated using OLS (estimated standard errors are given in parenthesis):



 

 

Perform a test of the null hypothesis that  in this model against the two sided alternative. What do you conclude? How would you interpret the OLS estimate of  in this equation? Interpret the value of the . *[20 marks]*

d) The model is extended to include the number of household members () and the age of the household head ().



The  obtained from estimating this model for the same 1,500 households using OLS is .

i) Perform a test between this and the regression model given in part c). *[10 marks]*

ii) What are the consequences for this analysis if household size and income are highly correlated? *[10 marks]*

e) What are the consequences for OLS estimation of the model given in part d) if the variance of the error term, , is given by . Outline one possible way of proceeding with the analysis. *[15 marks]*

d) What assumptions are required for the OLS estimator of the parameters of the model given in part d) to be unbiased and efficient? What do these assumptions mean in practice? Are you satisfied that these assumptions are valid in this case? *[15 marks]*

**Section B**

**Please answer TWO from Q4, Q5, Q6, Q7**

Question 1 (100 Marks) – Identification & Simultaneous Equations Models.

Part (a): (50 Marks)

Suppose we have data on how well a class of Leaving Cert students perform in their Higher Level Leaving Cert Geography exam. Consider the case where for some reason, we are not able to access the scores on a few students – perhaps we have data from a questionnaire and 6 out of 120 students decline to answer. Let  denote result in percentage terms where students need at least 40 to pass and  denote whether we observe the result or not.

1. Consider the probability that a randomly selected student passes. If we assume that the data was missing at random, would the bounds for this probability be tighter than if no such assumption was placed on the data? Why or why not? Is this assumption refutable?
2. Assume there is no missing data, i.e. we have the distribution of marks for all 120 students in the class. Let  be a dummy variable equal to one if a student has travelled abroad before and zero otherwise. Suppose upon discovering that  and  an examiner states the following:

"Data indicate that travelling abroad increases the probability a student will score highly in geography. The effect of foreign travel is to increase the probability of scoring at least  from 0.1 to 0.5."

Does this statement accurately describe the empirical findings? Explain.

Part (b): (50 Marks)

Consider the following macroeconomic model in structural form:



1. What type of relationship does the investment equation describe? Explain your answer.
2. Check the identifiability of the investment equation. Show your workings. Based on your answer, motivate an appropriate estimation technique for the investment equation.

Question 2 (100 Marks) – Limited Dependent Variables & Instrument Variables.

Part (a): (50 Marks)

Let  be the number of extramarital affairs for a married woman from the US population. The goal is to explain this variable (  ) in terms of other characteristics of the woman, especially whether she works outside of the home, her husband and her family. Is this a good candidate for a Tobit model? Explain. Now consider



Assume that the latent variable  satisfies Classical Linear Regression Model assumptions, in particular it is Normally distributed with homoscedastic variance and a linear conditional mean. How do we estimate such a Tobit model?

Distinguish a truncated model from a censored model. Under what circumstances might one entertain the possibility of employing a censored model and how do we estimate such models? How might we estimate a truncated model and if we used OLS, how might you expect the estimates to be affected by truncation?

Part (b): (50 Marks)

Suppose that you are interested in estimating the following expanded wage regression model

 (2)

where  is years of experience and we assume that  and  are uncorrelated with . Why might we include the quadratic term ? Let us suppose that you have two excluded exogenous regressors (instruments)  and  where  is the years of mother's education and  is years of father's education. Why might we want to test for endogeneity?

Estimating equation (2) by 2SLS in Stata, we obtain the following results

log(wage) = .048 + .061educ + .044exper - .0009exper2

(.400) (.031) (.013) (.0004)

*n* = 428, *R*2 = .136

Interpret the coefficient on  . Why might the 2SLS estimate be barely significant at the  level against a two-sided alternative? Is  a cause for concern? Explain.

© UNIVERSITY OF DUBLIN 2013