### **CARDIFF UNIVERSITY**

**examination paper**

**Academic Year: 2015/2016**

**Examination Period AUTUMN**

**Examination Paper Number: BST169**

**Examination Paper Title: ECONOMETRICS**

**Duration: 3 HOURS**

**Do not turn this page over until instructed to do so by the Senior Invigilator**

**Structure of Examination Paper:**

There are 4 pages.

There are 6 questions in total.

There are no appendices

The maximum mark for the examination paper is 100%. The percentage for each question is given in parentheses.

**Students to be provided with:**

Answer book

**Instructions to Students:**

All questions must be answered.

The use of non-electronic translation dictionaries between English or Welsh and a foreign language bearing an appropriate departmental stamp is permitted in this examination.

*Please turn over*

1. For the linear regression model:

y= X1β1 +x2β2+ε [1.1]

where y = (y1, y2, . . . , yN)’, X1 is an N × K matrix of explanatory variables with ones in the first column, x2 is an N ×1 vector, E(ε|X1, x2)=0 and Var(ε|X1, x2)=σ2IN.

* 1. Express the OLS estimator for β2 in [1.1] in terms of y, X1 and x2, and its variance in terms of σ2, X1 and x2. (4%)
  2. If ε follows a normal distribution, under H0: β2 =0 show that

[1.2]

where . (3%)

* 1. What distribution will the expression in [1.2] follow if σ2 is replaced by ? Why? (6%)
  2. If ε is not normally distributed, would your answer for the previous question be the same when the sample size tends to infinity? What assumptions are required? (5%)

1. {yi} is a sequence of i.i.d. random variables with density function p(yi|θ) for i=1,2,...N, where θ is a p by 1 vector. Answer the following questions.

a) What are the regularity conditions required for the density function such that  and ? (6%)

b) Suppose R is a q by p matrix with rank equal to q, which is less than p. For H0: Rθ=0, show that the likelihood ratio test statistic is equivalent to the Wald test statistic asymptotically and find the asymptotic distribution they follow. (10%)

1. Suppose yi, x1i, u2i and u3i are all random scalars following normal distributions and are mutually uncorrelated. Each variable is independent over i for i=1, 2,…, N with , and , which are assumed to be known. Additionally, x2i and x3i are generated from the following processes:

x2i= a0+a1x1i+ a2yi +u2i, [3.1]

x3i= a0+a1x2i+ a2yi +u3i, [3.2]

1. If yi is omitted in [3.1], will the OLS estimator for a1 be consistent? What about the OLS estimator for a1 in [3.2] if yi is omitted? Why? (4%)
2. Calculate the asymptotic bias for the OLS estimators in a). (4%)
3. Find the conditional covariance of yi and x1i given x2i, i.e. Cov(yi, x1i | x2i). (4%)
4. Consider the following regression,

yi=b0+b1x1i+b2x2i+εi. [3.3]

Express the probability limits of the OLS estimator for b1 in [3.3] in terms of the coefficients in [3.1] and the variances of yi, x1i, and u2i. Would it be 0 if both a1 and a2 are not 0? What is the implication herein for empirical studies? (6%)

1. Suppose {Zi} is a sequence of independent random variables, which are NOT from the same distribution. Answer the following questions.

a) What are the conditions for  to converge almost surely to , where μi=E(Zi)? (6%)

b) Can you use the results from a) to state the conditions under which the OLS estimators for β1 below is consistent, where yi and xi are scalars with i=1, 2… N? Prove it. (10%)

yi=β0+xi β1+εi

1. Consider the following two models to study spread betting in basketball games,

P(favwin=1|spread)= β0+ β1 spread [5.1]

P(favwin=1|spread)= Φ(α0+ α1 spread) [5.2]

where favwin is a binary variable: 1 for the favoured team in the game to win and 0 otherwise; spread denotes the score spread given by the gambling company before the game is played. Φ(ˑ) is the cumulative distribution function of standard normal distribution.

1. What values do you expect β0 in [5.1] and α0 in [5.2] to take if spread incorporates all the relevant information about the game outcome? (4%)
2. The estimation results of [5.1] from 500 observations are shown below,

P(favwin=1|spread)= 0.5769 +0.01937spread

(0.028) (0.0023)

[0.052] [0.0019]

where the numbers inside the parentheses (ˑ) are standard errors unadjusted for heteroskedasticity and the numbers inside the squared brackets [ˑ] are heteroskedasticity consistent standard errors.

Test whether spread incorporates all the relevant information about the game’s outcome in this model at 5% level of significance. (4%)

1. The estimation results of [5.2] are shown below  
   P(favwin=1|spread)= Φ(-0.010593+ 0.092463 spread),  
    (0.1037) (0.0122)   
   where the standard errors are shown in parentheses and the log likelihood function evaluated at the maximum is -263.5622.  
   Is the probability for the favoured team to win when spread=5 bigger than 50%? Why? (4%)
2. One now adds more explanatory variables, favhome (favoured team at home), fav25 (the ranking of the favoured team is in top 25) and und25 (underdog in top25), into [5.2] and obtains the following estimation results,

P(favwin=1|spread)= Φ(-0.055+ 0.088spread+0.15favhome+0.003fav25-0.22und25)

where the log likelihood function evaluated at the maximum is -262.6418. Test whether the additional explanatory variables are jointly significantly different from 0 at 5% level of significance. (4%)

Some potentially useful quantile function values from R are:

qchisq(0.95,3)=7.81, qf(0.95,3,498)=2.62, qt(0.975,498)=1.96.

1. Given the data matrix W=[y,ι,x,z], where y, x and z are column vectors with all the observations stacked up and ι is a vector of ones, one knows:

.

For the linear regression model,

y=ιβ0+x β1+ε

Var(ε|x,z)=σ2I, where I is the identity matrix.

1. Find the sample size and the ordinary least squares (OLS) estimates for β0 and β1. (5%)
2. Suppose x is endogenous and z is a valid instrument variable (IV), how would you estimate β0 and β1 with z? What is the likely sign of the correlation between x and ε? (5%)
3. Under IV estimation, how would you test whether β1 is significantly different from 0? What is your conclusion? (6%)

Some potentially useful quantile function values from R are:

qt(0.975,18)=2.1, qnorm(0.975)=1.96