BIRKBECK
(University of London)

MSc EXAMINATION FOR INTERNAL STUDENTS

Department of Economics, Mathematics and Statistics

ECONOMETRICS

EMEC026S7

Friday, 8 January 2021, 10.00 am - 2.00 pm (includes 2 hours for scanning and uploading)

The paper is divided into two sections. There are three questions in section A and three questions in section B.

Answer **ONE** question from SECTION A, and **ONE** question from SECTION B, and **ONE** further question from either section. All questions carry the same weight; the relative weight of sub-questions is indicated in square brackets.

Non-programmable calculators are allowed.

In some questions, you may require 5% critical values for the Chi-squared distribution, to carry out tests. These critical values, CV, for various degrees of freedom, DoF, are:

DoF	1	2	3	4	5	6	7	8	9
CV	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92

Section A

1. In the model

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u},$$

$\mathbf{y}$ is a $T \times 1$ vector of observations on a dependent variable, $\mathbf{X}$ is a $T \times k$ full-rank matrix of observations on regressors, β is a $k \times 1$ vector of unknown parameters, and $\mathbf{u}$ is a $T \times 1$ vector of unobserved disturbances. For each of the three cases (a)-(c) listed below: (1) derive an estimator of β; (2) derive the variance-covariance matrix for that estimator; (3) explain how you would estimate σ^2.

(a) [25%] $E(\mathbf{u}) = 0$, $E(\mathbf{uu}') = \sigma^2\mathbf{I}$, and $E(\mathbf{X}'\mathbf{u}) = 0$;

(b) [25%] $E(\mathbf{u}) = 0$, $E(\mathbf{uu}') = \sigma^2\mathbf{\Omega}$, where $\mathbf{\Omega}$ is a known positive definite symmetric matrix and $E(\mathbf{X}'\mathbf{u}) = 0$;

(c) [25%] $E(\mathbf{u}) = 0$, $E(\mathbf{uu}') = \sigma^2\mathbf{I}$, and $E(\mathbf{X}'\mathbf{u}) \neq 0$, but there exists a $T \times k$ matrix of instruments $\mathbf{W}$ such that $E(\mathbf{W}'\mathbf{u}) = \mathbf{0}$ and $E(\mathbf{W}'\mathbf{X}) \neq \mathbf{0}$.

(d) [25%] Suppose that β is estimated by ordinary least squares, OLS, and normality of the residuals is rejected because of severe excess kurtosis, but the Gauss-Markov assumptions, GMA, hold. Consider the argument that, since normality is not one of the GM assumptions, OLS is minimum variance in the class of linear unbiased estimators, thus there is nothing to be gained by estimating by maximum likelihood using a fat tailed distribution like Student's t.

2. Consider the linear regression model

$$y_t = \beta' x_t + u_t, \qquad t = 1, 2, ..., T,$$

where y_t is an observation on a dependent variable at time t; x_t is a $k \times 1$ vector of observations on some exogenous regressors; u_t is an unobserved disturbance, and β is a $k \times 1$ vector of unknown coefficients. For each of the 'problems' listed below: (i) explain the consequences of the problem for the properties of the least squares estimates of β and their standard errors; (ii) explain how you would test for the presence of the problem (iii) briefly indicate how you might resolve the problem if it is detected.

(a) [20%] Fourth-order serial correlation of the disturbance u_t;

(b) [20%] The variance of u_t was not constant, but related to x_t;

(c) [20%] Incorrect functional form;

(d) [20%] A shift in the variance of the errors at a known time $T1$, with $k < T1 < T - k$;

(e) [20%] A shift in the coefficients at a known time $T1$, with $k < T1 < T - k$.

3. Consider a third order autoregressive process, where ε_t is an independent normal white noise process, $\varepsilon_t \sim IN(0, \sigma^2)$, and there is a sample of data for y_t, $t = 1, 2, ..., T$,

$$y_t = \alpha + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \rho_3 y_{t-3} + \varepsilon_t. \tag{1}$$

   (a) [20%] What is meant by a covariance stationary stochastic process? What is meant by a stochastic process being integrated or order zero, $I(0)$, one, $I(1)$, and two, $I(2)$.

   (b) [20%] How would you rewrite (1) to test H_0 that y_t is $I(1)$ against H_1 that it is $I(0)$? Does this impose any restrictions on (1)? What is the test statistic?

   (c) [20%] How would you rewrite (1) to test H_0 that y_t is $I(2)$ against H_1 that it is $I(1)$? Does this impose any restrictions on (1)? What is the test statistic?

   (d) [20%] Suppose that (1) included a time trend, say γt. How would you then rewrite it to test H_0 that y_t is $I(1)$ against H_1 that it is $I(0)$ after removal of a deterministic trend. What restriction is required to ensure that there is not a quadratic trend in y_t under H_0.

   (e) [20%] Explain why it can be difficult in practice to determine the order of integration.

Section B

4. Quarterly UK data 1970Q3-2016Q4 is used to estimate a partial adjustment process: $\Delta r_t = \lambda(r_t^* - r_{t-1}) + \varepsilon_t$, to a long run "Taylor Rule" : $r_t^* = \theta_0 + \theta_1 \Delta q_t + \theta_2 \Delta p_t$. The nominal short-term interest rate per quarter, r_t, is computed as $0.25 \ln(1 + R_t/100)$ where R_t is the annual percentage rate, Δq_t is the change in log real income, and Δp_t is the change in log price level. The first equation, which is estimated by non-linear least squares, gives coefficients (standard errors) of

$$\begin{array}{llllllll} \Delta r_t = & 0.08(& -0.01 & +1.45 & \Delta q_t & +1.66 & \Delta p_t & -r_{t-1}) \quad +u_{1t} \\ & (0.015) & (0.004) & (0.33) & & (0.24) & & \end{array}$$

The second equation, which is estimated by instrumental variables, treats Δq_t and Δp_t as endogenous and uses Δq_{t-1}, Δp_{t-1}, Δq_t^*, Δp_t^* as additional instruments, where Δq_t^*, Δp_t^* are changes in the logs of world real income and prices, is

$$\begin{array}{lllllllll} r_t = & -0.001 & +0.176 & \Delta q_t & +0.160 & \Delta p_t & +0.910 & r_{t-1} & +u_{2t} \\ & (0.0003) & (0.0308) & & (0.0304) & & (0.0199) & & \end{array}$$

This has a Sargan-Hansen $\chi^2(2)$ test statistic of 0.73. The reduced form equations for Δq_t and Δp_t were estimated and the residuals saved as $\widetilde{\Delta q}_t$ and $\widetilde{\Delta p}_t$. The F statistic in the Δq_t regression was 18 and in the Δp_t regression was 74. The final regression estimated by least squares is

$$\begin{array}{llllllll} r_t = & -0.001 & +0.176 & \Delta q_t & +0.160 & \Delta p_t & +0.910 & r_{t-1} \\ & (0.0003) & (0.0292) & & (0.0288) & & (0.0189) & \end{array}$$

$$\begin{array}{lllll} -0.10 & \widetilde{\Delta q}_t & -0.04 & \widetilde{\Delta p}_t & +u_{3t} \\ (0.0369) & & (0.0399) & & \end{array}$$

(a) [25%] Explain the relationship between the first and second specifications. Given the measurement units for the variables do you think the long-run Taylor rule coefficients θ_i, $i = 0, 1, 2$ in the two equations are plausible?

(b) [25%] Explain how many over-identifying restrictions there are in the second equation. Are they rejected?

(c) [25%] What are the explanatory variables in the reduced form regressions? Is there any evidence of a weak instrument problem?

(d) [25%] Explain the relationship between the second and third equations. What can you conclude about the endogeneity of growth and inflation from the third equation.

Please turn over

5. You are asked to advise on the specification of an equation explaining log earnings on UK data. There is a large sample of workers with data on their earnings, age, gender (male, female), years of education and highest level qualification (GCSE, A Level, BSc, Postgraduate, PhD) and which of the 12 regions of the UK they were employed in.

   (a) [20%] Explain the "dummy variable trap". Why might it arise here? How would you deal with it?

   (b) [20%] It is believed that earnings rise and then fall with age and that the age at which earnings peak is later for those with more years of education. How would you model this and estimate peak earnings age as a function of years of education.

   (c) [20%] There is a dispute as to whether education is better measured by the continuous variable years or the dummy variables for the highest qualification. What are the advantages and disadvantages of each and explain how would you try to resolve the dispute?

   (d) [20%] There is a dispute as to whether the male and female data should be pooled with a dummy for males or whether separate equations should be estimated for each. How would you try to resolve the dispute?

   (e) [20%] In the light of your answers above, set out the equation(s) that you would estimate.

6. UK data, t =1950-2016, on the unemployment rate in percent, u_t, and log real GDP, q_t were used to estimate the "Okun's Law" type equations below, where $\Delta u_t = u_t - u_{t-1}$; $\Delta q_t = q_t - q_{t-1}$. The estimated coefficients (standard errors) are

$$
\begin{array}{rllllllll}
u_t = & \underset{(8.85)}{-12} & \underset{(0.09)}{+1.39} & u_{t-1} & \underset{(0.09)}{-0.42} & u_{t-2} & \underset{(2.90)}{-23} & q_t & \\
 & \underset{(5.40)}{+17} & q_{t-1} & \underset{(3.86)}{+8} & q_{t-2} & \underset{(0.03)}{-0.06} & t & +\varepsilon_{1t} & (2)
\end{array}
$$

$$
\begin{array}{ll}
R^2 = 0.983, \quad MLL = -30.58 & AIC = -37.58 \\
SER = 0.404 & BIC = -45.30
\end{array}
$$

$$
\begin{array}{rllllllll}
\Delta u_t = & \underset{(8.85)}{-12} & \underset{(0.02)}{-0.03} & u_{t-1} & \underset{(0.09)}{-0.42} & \Delta u_{t-1} & \underset{(2.90)}{-23} & \Delta q_t & \\
 & \underset{(1.10)}{+2} & q_{t-1} & \underset{(3.86)}{-8} & \Delta q_{t-1} & \underset{(0.03)}{-0.06} & t & +\varepsilon_{2t} & (3)
\end{array}
$$

$$
\begin{array}{ll}
R^2 = 0.756, \quad MLL = -30.58 & AIC = -37.58 \\
SER = 0.404 & BIC = -45.30
\end{array}
$$

$$
\Delta u_t = \underset{(0.63)}{3.00} \; \underset{(0.02)}{+0.51} \; \Delta u_{t-1} \; \underset{(2.64)}{-26} \; \Delta q_t \; \underset{(0.003)}{-0.01} \; t \; +\varepsilon_{3t} \qquad (4)
$$

$$
\begin{array}{ll}
R^2 = 0.727, \quad MLL = -34 & AIC = -38.34 \\
SER = 0.417 & BIC = -42.75
\end{array}
$$

$$
\Delta u_t = \underset{(0.09)}{0.62} \; \underset{(0.07)}{+0.55} \; \Delta u_{t-1} \; \underset{(2.81)}{-24} \; \Delta q_t \; +\varepsilon_{4t} \qquad (5)
$$

$$
\begin{array}{ll}
R^2 = 0.665, \quad MLL = -41.15 & AIC = -44.15 \\
SER = 0.457 & BIC = -47.46
\end{array}
$$

where SER is the standard error of regression; MLL is the maximised log likelihood; $AIC = MLL_i - k_i$, is the Akaike Information Criterion; and $BIC = MLL_i - 0.5k_i \ln T$ is the Schwarz Bayesian Information Criterion , where k_i is the number of parameters in model i.

Questions are on the next page

(a) [20%] Show what the relationship is between the parameters in model (2) and model (3)? What is the long-run effect of log output on unemployment in (3)? Comment on your estimate.

(b) [20%] Consider the hypothesis that the coefficients of both u_{t-1} and q_{t-1} in (3) are equal to zero. Interpret the hypothesis. The Likelihood Ratio χ^2 test statistic for the hypothesis is 4.239. Is the hypothesis rejected using the standard critical value? How would you respond to the argument that your conclusion is not valid since the distribution of the test statistic is non-standard.

(c) [20%] What restrictions on model (2) would lead to model (4)? Interpret the coefficient of t in model (4). Does this seem plausible?

(d) [20%] In model (5) at what growth rate would unemployment be constant? Does this seem plausible?

(e) [20%] Which model would be chosen by the AIC? Which model would be chosen by the BIC? Which model would you choose? Explain your criteria.

End of Examination Paper