ECONOMICS TRIPOS Part IIB

Thursday 4 June 2009 9 to 12

Paper 10

THEORY AND PRACTICE OF ECONOMETRICS II

*You should answer **two** questions from Section A and **one** question from Section B. All questions carry equal weight.*

*Answers from **each** Section **must** be written in a separate booklet.*

*Write your number **not** your name on the cover sheet of **each** booklet*

STATIONERY REQUIREMENTS	SPECIAL REQUIREMENTS
20 Page booklet x 2	*Durbin Watson and Dickey Fuller and other tables*
Metric Graph Paper	*New Cambridge Elementary Statistical Tables*
Rough work pads	*Approved calculators allowed*
Tags	

You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator

SECTION A

1. Historically the average ratio of the price of gold to the price of silver has been about 50 though it has fluctuated well away from this. An investor is interested in whether this might represent a stable equilibrium that could be used for an investment strategy in gold, by selling gold when its price exceeds 50 times the price of silver, and buying gold when its price falls below 50 times the price of silver. To assist the investor an econometrician collects annual data for 1918-2008 on the average London fix for the dollar price of gold and silver. The variables are defined as follows:

   - LGOLD = ln(price of gold in dollars)
   - LSILVER = ln(price of silver in dollars)
   - RATIO = LGOLD-LSILVER
   - *trend*= time trend that is 1 in 1918, 2 in 1919 etc.
   - DUM79 is a dummy that is one in 1979 and zero elsewhere. [In 1979 former billionaire Nelson Bunker Hunt and his brother William Herbert Hunt attempted to corner the global silver market leading to an extraordinary spike in the price of silver. Nelson Bunker Hunt filed for bankruptcy in September 1988, largely due to lawsuits incurred as a result of his silver speculation].

   The econometrician estimates the following over this sample period (estimated coefficient standard errors are in parentheses, Δ is the first difference operator):

- Equation 1

$$LGOLD_t = \underset{(0.051)}{3.8221} + \underset{(0.041)}{1.0198} LSILVER_t - \underset{(0.451)}{1.2388} DUM79 + u_t \quad (Eq1a)$$

$$\widehat{\Delta u_t} = \underset{(0.045)}{-0.1371} u_{t-1} \quad (Eq1b)$$

$$\widehat{\Delta u_t} = \underset{(0.045)}{-0.1178} u_{t-1} - \underset{(0.100)}{0.2328} \Delta u_{t-1} \quad (Eq1c)$$

- Equation 2

$$\Delta\widehat{RATIO}_t = \underset{(0.195)}{0.7057} - \underset{(0.048)}{0.1843} RATIO_{t-1} - \underset{(0.198)}{0.9305} DUM79 + \underset{(0.001)}{0.0002} trend \quad (Eq2a)$$

$$\Delta\widehat{RATIO}_t = \underset{(0.194)}{0.6399} - \underset{(0.049)}{0.1605} RATIO_{t-1} - \underset{(0.194)}{0.9144} DUM79 + \underset{(0.001)}{0.0000} trend - \underset{(0.092)}{0.1871} \Delta RATIO_{t-1} \quad (Eq2b)$$

- Equation 3

$$\begin{pmatrix} \Delta\widehat{LGOLD}_t \\ \Delta\widehat{LSILVER}_t \end{pmatrix} = \begin{pmatrix} 0.1131 \\ (0.120) \\ -0.4433 \\ (0.202) \end{pmatrix} + \begin{pmatrix} 0.1461 & 0.2943 \\ (0.095) & (0.060) \\ 0.0809 & 0.0600 \\ (0.161) & (0.102) \end{pmatrix} \begin{pmatrix} \Delta LGOLD_{t-1} \\ \Delta LSILVER_{t-1} \end{pmatrix} - \begin{pmatrix} 0.0232 \\ (0.031) \\ -0.1179 \\ (0.052) \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} LGOLD_{t-1} \\ LSILVER_{t-1} \end{pmatrix} + \begin{pmatrix} 0.3226 \\ (0.126) \\ 1.2983 \\ (0.212) \end{pmatrix} DUM79 \quad \begin{matrix} (Eq3a) \\ (Eq3b) \end{matrix}$$

Selected diagnostics for the above equations are tabulated below:

	Eq1a	Eq1b	Eq1c	Eq2a	Eq2b	Eq3a	Eq3b
Sample	1918-2008	1920-2008					
Obs	91	89	89	89	89	89	89
R^2	0.877	0.09	0.14	0.288	0.321	0.41	0.34
se	0.436	0.183	0.178	0.196	0.192	0.12	0.21
DW	0.20	2.17	1.68	1.78	1.51	1.64	1.55
SBC	1.292	-0.523	-0.533	-0.269	-0.266	-1.155	-0.110

where se is the equation standard error, DW is the Durbin Watson statistic, SBC is Schwarz's Bayesian Information Criterion defined as $-2logL/T + klogT/T$ where T =no. of Obs, k =no. of explanatory variables and $logL$ is the equation log likelihood.

Explain what can be inferred from each of these equations and use them to assess the investor's proposed strategy. You may assume that critical values for any statistical tests you carry out are unaffected by the presence of DUM79 in estimated equations.

2. An econometrician interested in the effect of a minimum wage on youth employment collects annual data on the panel of 48 mainland states of the US from 1976 to 2007. She estimates the following specification

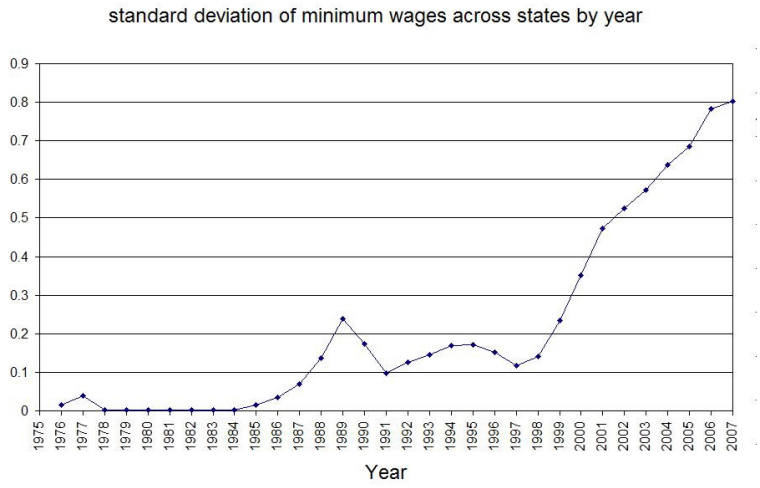
$$ln(empy_{it}) = \alpha_i + f_t + \gamma_i\, t + \beta ln(mw_{it}) + \varepsilon_{it}$$

where $empy_{it}$ is the youth employment-population ratio and mw_{it} the minimum wage in state i in year t, α_i and f_t are state- and time-specific fixed effects and t is a time trend.

She obtains the following results estimating this equation over different time periods

Sample Period	1976-1987	1976-1997	1976-2007
$\widehat{\beta}$	0.234	0.048	-0.134
s.e. $(\widehat{\beta})$	0.359	0.098	0.050
R^2	0.933	0.893	0.899

She also records the standard deviation across states of minimum wages over time as in the graph below



(a) Assess the evidence for a negative effect of minimum wages on youth employment over these sample periods. To what extent does the graph help explain these findings?

(b) What interpretation would you give to:

(i) the individual fixed effects?

(ii) the time fixed effects?

(iii) the state specific time trend coefficients?

(c) A colleague suggests that the regression is misspecified in that the explanatory variable should be real minimum wages (ie the right hand side should be $ln(mw) - ln(p)$ where p is a US-wide price index). Explain whether this criticism is valid.

(d) Another colleague argues that minimum wages may be endogenous in that state authorities raise minimum wages when the economy is strong. What effect would this have on your interpretation of these regression results and how might any inferential problems be overcome?

(e) A third colleague suggests that the effect of minimum wages on employment may be different in a recession. Explain how, given an indicator variable that is one if the US economy is in recession and zero otherwise, the above specification could be modified to investigate this hypothesis.

3. Comment on each of the following statements. In each case, state whether the statement is true or false, and explain the reasoning behind your answer.

   (a) The parameters of conventional macroeconomic relationships will be structurally unstable unless they are chosen to embody the underlying characteristics of preferences and technology. However this does not rule out using reduced form methods for forecasting and causality testing.

   (b) Even if macroeconomic time series are nonstationary, we can still estimate conventional levels regressions provided the series concerned are cointegrated. So there is nothing to be gained by expressing such relationships as Error Correction Mechanisms.

   (c) Even if we are concerned about omitted variable bias when working with cross-section data, we can always ensure that we obtain consistent parameter estimates by using instrumental variable techniques.

4. In a market comprised of J alternatives, on each purchase occasion consumer i allocates her income y_i on (at most) one of the j alternatives, and an "outside" good. The utility maximisation problem is given by

$$\max \; U_{ij}(\mathbf{v}_j, z) \;\text{ subject to }\; p_j + p_z z = y_i \qquad (1)$$

   where $\mathbf{v}_j$ are attributes of product j and p_j the price, z is the quantity of the outside good and p_z is its price. Consumer i's indirect utility function for product j is given by

$$\begin{aligned} U_{ij} &= \alpha_j + \varepsilon_{ij}, \quad j = 1, ..., J \\ \alpha_j &= \sum_{l=1}^{L} \omega_l v_{jl} - \omega_p v_p + \xi_j + \varepsilon_{ij} \end{aligned}$$

   where α_j, mean utility, is additive in a vector of *observed* product characteristics, and a product characteristic (ξ_j) that is observed by the consumer but unobserved by the econometrician, ε_{ij} is an error term denoting the deviation from the mean utility for individual i, and $\omega_l \quad l = 1, ..., L$ and ω_p are unknown parameters.

(a) Derive an expression for the aggregate demand (share) for product j assuming that consumer types are distributed type 1 extreme value.

(b) In what sense does the estimation of demand models in a market with product differentiation and in which a consumer buys at most one of a product give difficulties relative to traditional models of demand?

(c) Show that subsequent to a change in an attribute of a given product, the demand expression derived in (a) implies proportional substitution.

(d) You are told that price is endogeneous. Briefly evaluate the statement that the solution to endogeneity in nonlinear demand systems such as (1) is nonstandard. By locating the problem of endogeneity within a linear regression, demonstrate one solution to this problem.

(e) If we assume that the demand system is Nash in prices what type of instruments might you use? If there are no instruments available, discuss how brand fixed effects might be used to at least partially ameliorate the problem of endogeneity.

5. (a) Explain what you understand by the terms "average treatment effect" (ATE) and "average treatment effect on the treated" (ATET).

(b) Suppose that you have estimates of the ATE and ATET both conditional on a covariate vector x. How might estimates of the unconditional ATE and ATET be obtained?

(c) What are the Hawthorn and John Henry effects? How might their presence affect the validity of conclusions drawn in a randomised experimental setting?

(d) There are two groups, treatment $j = 1$ and control $j = 0$. There are two time periods $t = a$ and b. Treatment occurs when $t = a$ for the treatment group $j = 1$. Justify the use of the *difference in differences* estimator

$$(\bar{y}_{1,a} - \bar{y}_{1,b}) - (\bar{y}_{0,a} - \bar{y}_{0,b}),$$

as an estimator of ATE where $\bar{y}_{j,t}$ denotes the sample mean outcome for group $j = 0, 1$ at time $t = a, b$.

SECTION B

6. Business cycle data are often held to show asymmetries in both the duration and magnitude of expansions and contractions. Explain carefully how ARMA modelling can be used to characterise the behaviour of output growth and how such simple univariate time series techniques can be modified to account for such asymmetries.

7. Evaluate critically the statement below and discuss the econometric issues that have influenced the development of models of demand systems over differentiated products such that this condition is upheld.

   For a given market we let Ω_J denote a finite choice set of differentiated products. When we introduce a new product (say product "A") that is virtually identical to an existing product we would expect that the combined markets shares of the new product and product "A" to be approximately the same as was the market share of product "A" before we introduced the new product.

8. If agents' expectations of the future affect their behaviour today, then our ability to distinguish rival economic theories becomes almost nonexistent. Discuss giving empirical examples.

9. The econometric difficulties in estimating the factors that determine growth mean that in practice all we have learned from growth regressions is that we don't know very much about economic growth. Discuss.

END OF PAPER