UNIVERSITY OF VICTORIA

EXAMINATIONS, APRIL 2007

ECONOMICS 546: THEMES IN ECONOMETRICS

TO BE ANSWERED IN BOOKLETS **DURATION: 3 HOURS**
INSTRUCTOR: D. Giles

STUDENTS MUST COUNT THE NUMBER OF PAGES IN THIS EXAMINATION PAPER BEFORE BEGINNING TO WRITE, AND REPORT ANY DISCREPANCY IMMEDIATELY TO THE INVIGILATOR.

THIS QUESTION PAPER HAS 7 PAGES.
STATISTICAL TABLES ARE SUPPLIED SEPARATELY.

This is a "closed book/closed notes" examination.
Calculators may be used.
Answer ALL QUESTIONS (***Total Marks = 90***)

Question 1:

Critically discuss and appraise the following statement: "The Bayesian approach to estimation has all of the advantages, and none of the disadvantages, of Maximum Likelihood estimation. So, it is difficult to comprehend why some econometricians aren't Bayesians!"

Total: 6 marks

Question 2:

Suppose that we have a random sample of *n* observations from a Pareto distribution, with a ***known*** location parameter, y_m, and an ***unknown*** shape parameter, *k*. That is, the density function for an individual observation is:

$$p(y_i \mid k, y_m) = k\, y_m^k / [y_i^{k+1}] \quad ; \quad y_m < y_i < \infty \quad ; \quad k > 0 .$$

(a) As a Bayesian, suppose that I decide to represent my prior uncertainty about *k* with a prior density which is Gamma, with a shape parameter, α (> 0), and a scale parameter, θ (> 0). That is:

$$p(k) = k^{\alpha-1} e^{-k/\theta} / [\theta^{\alpha}\Gamma(\alpha)] \quad ; \qquad k > 0 .$$

$\Gamma(\alpha)$ is a Gamma function. (It is just a *constant* once we assign a value to α). The mean of this distribution is $(\alpha\, \theta)$, its variance is $(\alpha\, \theta^2)$. Its mode is at $[(\alpha\text{-}1)\theta]$, if $\alpha > 1$.

Show that the posterior density for k is also Gamma, but with a shape parameter which is $(n + \alpha)$, and a scale parameter which is $[\theta^{-1} + \sum_{i=1}^{n} \log_e(y_i / y_m)]^{-1}$.

8 marks

(b) What is the Bayes estimator of k if I have a quadratic loss function? What is the Bayes estimator of k if I have a zero-one loss function?

3 marks

(c) Show that the Maximum Likelihood Estimator (MLE) of k is $\tilde{k} = [n / \sum_{i=1}^{n} \log_e(y_i / y_m)]$.

4 marks

(d) Show that the Bayes estimator you obtained in part (b) above under quadratic loss collapses to the MLE for k if $\theta \to \infty$ and $\alpha \to 0$. Why does this result make sense intuitively? (What is happening to the prior density in this situation?)

2 marks
Total: 17 Marks

Question 3:

Consider the situation in Question 2, where our data follow a Pareto distribution. In that case, the MLE for the parameter, k, is $\tilde{k} = [n / \sum_{i=1}^{n} \log_e(y_i / y_m)]$. You may also want to note the following features of the Pareto distribution, using the notation in the opening statement of Question 2:

Its mean is $[ky_m / (k - 1)]$, if $k > 1$; its median is $[y_m 2^{1/k}]$; its mode is y_m; and its variance is $[(ky_m^2) / \{(k-1)^2 (k-2)\}]$, if $k > 2$.

(a) Construct the Likelihood Ratio Test statistic for testing H_0: $k = 2$ *vs*. H_A: $k \neq 2$. Apply the LRT using the values n = 100, y_m = 1, and $\sum_{i=1}^{n} \log_e(y_i)$ = 25. Explain what you conclude about the existence of the variance of the underlying distribution, and why.

9 marks

(b) Construct the Lagrange Multiplier test statistic for testing H_0: $k = 2$ *vs*. H_A: $k \neq 2$. Apply the LM test using the values n = 100, y_m = 1, and $\sum_{i=1}^{n} \log_e(y_i)$ = 25. Explain what you conclude about the existence of the mean of the underlying distribution, and why.

9 marks

(c) What is the MLE of the ratio, (median / mode), for this distribution? What “good” properties does this estimator possess?

3 marks
Total: 21 Marks

Question 4:

Suppose that we use a Poisson regression model to model some "count" data, y_1, y_2, y_3, ……..
So, we specify:

$$\Pr.[Y = y_i] = \exp(-\mu_i)\mu_i^{y_i} / y_i! \tag{1}$$

and then set

$$\mu_i = \exp[x_i'\beta] \qquad ; \quad i = 1, 2, 3, \ldots.., n \tag{2}$$

(a) In what sense is equation (2) analogous to what we do when we set up a conventional regression model? Why do we use the exponential function in equation (2)?

3 marks

(b) Write down the log-likelihood function for this model, based on n independent observations, and show that the first-order condition for maximizing it is:

$$\sum_{i=1}^{n}[y_i - \exp(x_i'\beta)]x_i = 0\,.$$

3 marks

(c) Why is it reasonable to refer to the quantity $[y_i - \exp(x_i'\tilde{\beta})]$ as the "i^{th} residual"?

2 marks

(d) What is the sum of these residuals if the covariates (the x variables) include an intercept?

3 marks

(e) Show that the marginal effect associated with the j^{th} covariate at observation i is $ME_{ij} = [\beta_j \exp(x_i'\beta)]$.

2 marks

(f) Prove that if we take the average of the j^{th} marginal effect over the n observations in the sample, and if an intercept is included in the model, then this ***average*** marginal effect is just $\tilde{\beta}_j \bar{y}$, where $\tilde{\beta}$ is the MLE, and $\bar{y}$ is the sample mean of the y-values.

3 marks
Total: 16 Marks

Question 5:

(a) ***Briefly*** explain what we mean by a "Seemingly Unrelated Regressions" (SUR) model, and why it may improve our ability to model certain situations.

4 marks

(b) ***State*** two conditions under which the SUR estimator will be identical to the OLS estimator for ***all*** of the coefficients in the model. (You don't have to prove these conditions.)

2 marks

(c) Suppose that we have a two-equation SUR model. It can be shown the SUR estimator for the coefficient vector in *just the first equation* will be identical to the OLS estimator for that vector if and only if $X_2(X_2'X_2)^{-1}X_2'X_1 = X_1$, where X_i denotes the regressor matrix in the i^{th} equation.

Suggest a simple situation that may arise with the regressor matrices that would ensure that this condition is satisfied.

4 marks

Total: 10 Marks

Question 6:

The "Almost Ideal Demand System" proposed by Deaton and Muellbauer comprises n demand equations, with the "budget shares" as the dependent variables, and takes the following form:

$$w_i = \alpha_i + \beta_i \log(M/P) + \sum_{j=1}^{n} \gamma_{ij} \log(p_j) + \varepsilon_i \quad ; \quad i = 1, 2, \ldots, n$$

where the i^{th} "budget share" is:

$$w_i \equiv (p_i q_i)/(\sum_{j=1}^{n} p_j q_j) \quad ; \quad i = 1, 2, \ldots, n$$

$$M \equiv \sum_{j=1}^{n} p_j q_j$$

is total expenditure (or "income"), and

$$\log(P) \equiv \sum_{i=1}^{n} \alpha_i \log(p_i) + 0.5 \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \log(p_i)\log(p_j)$$

is an aggregate price deflator.

(a) Engel aggregation implies that the following restrictions *must* be imposed on the parameters of the model in order for it to make any economic sense:

$$\sum_{i=1}^{n} \alpha_i = 1 \; ; \qquad \sum_{i=1}^{n} \beta_i = 0 \; ; \qquad \text{and} \qquad \sum_{i=1}^{n} \gamma_{ij} = 0 \;\; (\text{for } j = 1, 2, \ldots, n).$$

Briefly explain where these restrictions come from.

3 marks

(b) I have used annual data for alcohol consumption in the U.K. (1955 – 1985) to estimate a demand system of this type. There are *three* goods – beer, wine and spirits, but the model that I have estimated contains only *two* equations – the first for expenditure on beer, and the second for expenditure on wine. Why is this?

3 marks

The results of estimating this model using EViews are given in **Output 1** on the next page. You may wish to note that the way I have named the coefficients of the model is as follows:
C(1) = α_1; C(2) = α_2;
C(3) = β_1; C(4) = β_2;
C(5) = γ_{11}; C(6) = γ_{12}; C(7) = γ_{13}; C(8) = γ_{21}; C(9) = γ_{22}; C(10) = γ_{23}

Output 1: The first equation is for Beer; the second is for Wine

System: SYS01
Estimation Method: Seemingly Unrelated Regression
Date: 04/18/07 Time: 10:16
Sample: 1955 1985
Included observations: 31
Total system (balanced) observations 62
Iterate coefficients after one-step weighting matrix
Convergence achieved after: 1 weight matrix, 4 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.034387	0.027171	38.06956	0.0000
C(3)	2.78E-05	2.36E-05	1.180707	0.2431
C(2)	-0.054447	0.023981	-2.270430	0.0274
C(5)	0.041983	0.015595	2.692116	0.0095
C(6)	-0.078291	0.013958	-5.609211	0.0000
C(7)	0.012176	0.016178	0.752596	0.4551
C(8)	-0.031116	0.013768	-2.260110	0.0280
C(9)	0.063707	0.012322	5.170318	0.0000
C(10)	-0.008666	0.014277	-0.606953	0.5465
C(4)	-2.45E-05	2.08E-05	-1.177698	0.2443

Determinant residual covariance 1.91E-11

Equation: WB=C(1)+C(3)*(LM-C(1)*PB-C(2)*PW-(1-C(1)-C(2))*PS-0.5
*(C(5)*LPB*LPB+C(6)*LPB*LPW +C(7)*LPB*LPS+C(8)*LPW
*LPB+C(9)*LPW*LPW+C(10)*LPW*LPS+(0-C(5)-C(8))*LPS*LPB
+(0-C(6)-C(9))*LPS*LPW+(0-C(7)-C(10))*LPS*LPS))+C(5)*LPB
+C(6)*LPW+C(7)*LPS

Observations: 31

R-squared	0.946987	Mean dependent var	0.912493
Adjusted R-squared	0.927709	S.D. dependent var	0.018460
S.E. of regression	0.004963	Sum squared resid	0.000542
Durbin-Watson stat	1.004163		

Equation: WW=C(2)+C(4)*(LM-C(1)*PB-C(2)*PW-(1-C(1)-C(2))*PS-0.5
*(C(5)*LPB*LPB+C(6)*LPB*LPW +C(7)*LPB*LPS+C(8)*LPW
*LPB+C(9)*LPW*LPW+C(10)*LPW*LPS+(0-C(5)-C(8))*LPS*LPB
+(0-C(6)-C(9))*LPS*LPW+(0-C(7)-C(10))*LPS*LPS))+C(8)*LPB
+C(9)*LPW+C(10)*LPS

Observations: 31

R-squared	0.957844	Mean dependent var	0.065665
Adjusted R-squared	0.942515	S.D. dependent var	0.018276
S.E. of regression	0.004382	Sum squared resid	0.000422
Durbin-Watson stat	0.837702		

(c) What estimator has been used in **Output 1**? Is this an appropriate estimator to use here? In what way(s), if any, would the results differ if the OLS estimator were used?

3 marks

(d) It can be shown that the "income" elasticity of demand for the i^{th} good in the Almost Ideal Demand System is $(1 + \beta_i / w_i)$. Use the results in **Output 1** to compute a point estimate of the income elasticity of demand for ***spirits***.

4 marks

As you know, a demand system should be homogeneous of degree zero in prices and nominal income. In the case of the Almost Ideal Demand System, this means that the following *additional* restrictions on the parameters of the model must be satisfied:

$$\sum_{j=1}^{n} \gamma_{ij} = 0 \text{ (for } i = 1, 2, \ldots, n).$$

Output 2 shows the results of testing these restrictions (once the restrictions that have been imposed already to allow for Engel aggregation are taken into account):

Output 2:

Wald Test:
System: SYS01

Test Statistic	Value	df	Probability
Chi-square	18.52047	2	0.0001

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(5) + C(6) + C(7)	-0.024133	0.006488
C(8) + C(9) + C(10)	0.023925	0.005726

Restrictions are linear in coefficients.

(e) Interpret these test results.

2 marks

(f) **Output 3** on the next page shows the results of re-estimating the original model, but now imposing the additional parameter restrictions implied by homogeneity. Test the validity of the additional homogeneity restrictions, using a Likelihood Ratio Test.

[**Hint:** Recall that the log-likelihood function for the SUR model can be expressed as $\log L = const. - (T/2)\log|\tilde{\Sigma}|$.]

5 marks

Output 3:

System: SYS02
Estimation Method: Seemingly Unrelated Regression
Date: 04/18/07 Time: 10:23
Sample: 1955 1985
Included observations: 31
Total system (balanced) observations 62
Iterate coefficients after one-step weighting matrix
Convergence achieved after: 1 weight matrix, 3 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.933830	0.002570	363.4196	0.0000
C(3)	9.41E-05	2.00E-05	4.702347	0.0000
C(2)	0.045261	0.002359	19.18273	0.0000
C(5)	0.035494	0.018581	1.910301	0.0614
C(6)	-0.097671	0.015311	-6.379335	0.0000
C(8)	-0.024627	0.017061	-1.443432	0.1547
C(9)	0.082929	0.014060	5.898232	0.0000
C(4)	-9.00E-05	1.84E-05	-4.898554	0.0000

Determinant residual covariance 3.04E-11

Equation: WB=C(1)+C(3)*(LM-C(1)*PB-C(2)*PW-(1-C(1)-C(2))*PS-0.5
*(C(5)*LPB*LPB+C(6)*LPB*LPW +(0-C(5)-C(6))*LPB*LPS+C(8)
*LPW*LPB+C(9)*LPW*LPW+(0-C(8)-C(9))*LPW*LPS+(0-C(5)
-C(8))*LPS*LPB +(0-C(6)-C(9))*LPS*LPW+(0-(0-C(5)-C(6))-(0-C(8)
-C(9)))*LPS*LPS))+C(5)*LPB+C(6)*LPW+(0-C(5)-C(6))*LPS
Observations: 31

R-squared	0.923591	Mean dependent var	0.912493
Adjusted R-squared	0.904489	S.D. dependent var	0.018460
S.E. of regression	0.005705	Sum squared resid	0.000781
Durbin-Watson stat	1.045106		

Equation: WW=C(2)+C(4)*(LM-C(1)*PB-C(2)*PW-(1-C(1)-C(2))*PS-0.5
*(C(5)*LPB*LPB+C(6)*LPB*LPW +(0-C(5)-C(6))*LPB*LPS+C(8)
*LPW*LPB+C(9)*LPW*LPW+(0-C(8)-C(9))*LPW*LPS+(0-C(5)
-C(8))*LPS*LPB +(0-C(6)-C(9))*LPS*LPW+(0-(0-C(5)-C(6))-(0-C(8)
-C(9)))*LPS*LPS))+C(8)*LPB+C(9)*LPW+(0-C(8)-C(9))*LPS
Observations: 31

R-squared	0.934269	Mean dependent var	0.065665
Adjusted R-squared	0.917836	S.D. dependent var	0.018276
S.E. of regression	0.005239	Sum squared resid	0.000659
Durbin-Watson stat	0.821969		

Total: 20 Marks

END OF EXAMINATION