UNIVERSITY OF VICTORIA **EXAMINATIONS, APRIL 2010**

<u>ECONOMICS 546: THEMES IN ECONOMETRICS</u>

TO BE ANSWERED IN BOOKLETS **DURATION: <u>3 HOURS</u>**
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 <u>6 PAGES</u>.
STATISTICAL TABLES AND A FORMULA SHEET ARE SUPPLIED SEPARATELY.

This is a “closed book/closed notes” examination. Calculators may be used.
Answer <u>ALL FIVE QUESTIONS</u> **(*Total Marks = 90*)**

Question 1:

Discuss the following statement:
“When estimating the equations of a simultaneous equations model we have to trade off the advantages and disadvantages associated with ‘single equation’ estimators, and ‘full system’ estimators.”

Total: 10 marks

Question 2:

The probability density function for the Erlang distribution is:

$$p(x_i) = [(x_i / b)^{c-1} \exp(-x_i / b)]/[b(c-1)!] \quad ; \quad b, c > 0 \; ; \; 0 < x_i < \infty \; ; \; i = 1, 2, \ldots\ldots, n.$$

The scale parameter is b, and the shape parameter is c. For this distribution, the characteristic function is:

$$\phi_x(t) = (1 - itb)^{-c} \; ; \text{ where } i^2 = -1.$$

(a) Prove that the mean of this distribution is cb and its variance is cb^2

(5 marks)

(b) Now assume that the value of c is <u>known</u>, so the only parameter to be estimated is b. Derive the MLE for b. (Make sure that you check the second-order condition.)

(6 marks)

(c) Derive the Lagrange Multiplier Test of H_0: $b = 1$, against a 2-sided alternative hypothesis. Explain how you would apply this test in practice.

(6 marks)

(d) Derive the Wald Test of H_0: $b = 1$, against a 2-sided alternative hypothesis. Explain how you would apply this test in practice.

(3 marks)
Total: 20 marks

Question 3:

(a) Explain why the usual linear regression model is generally inappropriate when we wish to model "count" data.

(5 marks)

(b) An alternative distribution for count data is the Geometric distribution, whose mass function is :

$$p(y \mid \mu) = \mu^{y}(1+\mu)^{-(y+1)} \qquad ; \quad \mu > 0 \quad ; \quad y = 0, 1, 2, \ldots\ldots\ldots$$

Show that this distribution is a member of the "linear exponential family" of distributions – that is, show that its mass function can be written in the form:

$$p(y \mid \mu) = \exp[a(\mu) + b(y) + yc(\mu)] \,,$$

where $a(\mu)$ and $c(\mu)$ are functions only of μ ; and $b(y)$ is a function only of y.

[**Hint:** Write $p(y \mid \mu) = \exp[\log\{\mu^{y}(1+\mu)^{-(y+1)}\}]$, and proceed from there to show that $a(\mu) = -\log(1+\mu)$ and $c(\mu) = \log[\mu/(1+\mu)]$.]

(4 marks)

(c) The mean and variance of distributions in the linear exponential family are

$$E(Y) = -a'(\mu)/c'(\mu) \text{ and } Var(Y) = 1/c'(\mu) \text{ respectively, where } a'(\mu) = \frac{da(\mu)}{d\mu}, \textit{ etc.}$$

Show that the mean and variance of the Geometric distribution are μ and [μ (μ+1)] respectively.

(4 marks)

(d) What advantage might this distribution have over the Poisson distribution when modeling count data?

(2 marks)

(e) Now introduce "covariates" (explanatory variables) into the Geometric model by setting $\mu_i = \exp[x_i'\beta]$, where x_i and β are each $(k \times 1)$ non-random vectors. Let y_i be the i^{th} observation on y. Why don't we just let $\mu_i = (x_i'\beta)$?

Write down the likelihood function for this model, assuming n independent sample values, and show that the (k) "likelihood equations" are:

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

(10 marks)

Total: 25 marks

Question 4:

> It can be shown that if a Bayes estimator has **constant risk** (for all values of the parameters), then this Bayes estimator is "**mini-max**".

Suppose that the likelihood function is a Binomial distribution, so that

$$p(y \mid \theta, n) \propto \theta^{y}(1-\theta)^{n-y} \;; \qquad y = 0, 1, 2, \ldots, n \;; \qquad 0 < \theta < 1 .$$

In this case, the Natural Conjugate prior is a Beta distribution, with **positive** parameters a and b. The posterior for θ is then a Beta distribution as well, with parameters $(y + a)$ and $(n - y + b)$, so the mean of the posterior (the Bayes estimator under quadratic loss) is $\hat{\theta} = [(y + a) / (a + b + n)]$.

(a) Recalling that the mean of y is $(n\theta)$, and its variance is $(n\theta)(1 - \theta)$, show that:

(i) The bias of $\hat{\theta}$ is $[a - \theta(a + b)] / [a + b + n]$.

(ii) The variance of $\hat{\theta}$ is $[n\theta(1 - \theta)] / [(a + b + n)^2]$

(4 marks)

(b) Using these results, write down the risk of $\hat{\theta}$, under quadratic loss (*i.e.*, its MSE).

(2 marks)

(c) Show that the conditions under which this risk is **constant** (*i.e.*, independent of the value of θ) are $[(a + b)^2 - n] = 0$, and $[n - 2a(a + b)] = 0$.

(3 marks)

(d) Show that the Bayes estimator, $\hat{\theta}$, will be **mini-max** if the prior is chosen to be Beta with $a = b = (n / 4)^{½}$. Finally, draw a diagram of this risk function for $\hat{\theta}$ when $n = 1$.

(6 marks)

Total : 15 marks

Question 5:

I have estimated two regression equations, one explaining hourly earnings (wages), HREARN, and the other explaining hourly "benefits", HRBENS. Both HREARN and HRBENS are measured in dollars per hour. "Benefits" include things such the employer's contributions to medical and dental insurance, a pension plan, *etc*. The explanatory variables in my model are:

EDUC = years of schooling
EXPER = years of work experience
UNION = 1 if person is a union member, = 0 if not a member
MARRIED = 1 if person is married; = 0 if they are not married
WHITE = 1 if the person is white; = 0 if they are not white
MALE = 1 if the person is a male; = 0 if they are female

I have a sample of data for 616 individual employees in the same year.

(a) Consider OUTPUT 1 and OUTPUT 2 on the next 2 pages. Why are the estimated coefficients identical in the two outputs? What is different across the two outputs – why ?

(3 marks)

OUTPUT 1

System: SYS01
Estimation Method: Least Squares
Date: 13/04/10 Time: 13:21
Sample: 1 616
Included observations: 616
Total system (balanced) observations 1232

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-3.708936	1.070565	-3.464466	0.0005
C(2)	0.438941	0.067589	6.494293	0.0000
C(3)	0.099272	0.015586	6.369494	0.0000
C(4)	0.673529	0.394460	1.707469	0.0880
C(5)	0.454913	0.415003	1.096170	0.2732
C(6)	1.032313	0.613419	1.682884	0.0927
C(7)	1.770192	0.399066	4.435835	0.0000
C(11)	-0.752465	0.135846	-5.539098	0.0000
C(12)	0.079526	0.008576	9.272577	0.0000
C(13)	0.010248	0.001978	5.181783	0.0000
C(14)	0.464771	0.050054	9.285426	0.0000
C(15)	0.125161	0.052660	2.376753	0.0176
C(16)	0.102897	0.077838	1.321943	0.1864
C(17)	0.232670	0.050638	4.594744	0.0000
Determinant residual covariance		5.423959		

Equation: HREARN=C(1)+C(2)*EDUC+C(3)*EXPER+C(4)*UNION+C(5)
*MARRIED+C(6)*WHITE+C(7)*MALE
Observations: 616

R-squared	0.184373	Mean dependent var	6.229286
Adjusted R-squared	0.176337	S.D. dependent var	4.836833
S.E. of regression	4.389710	Sum squared resid	11735.16
Durbin-Watson stat	1.950503		

Equation: HRBENS=C(11)+C(12)*EDUC+C(13)*EXPER+C(14)*UNION
+C(15)*MARRIED+C(16)*WHITE+C(17)*MALE
Observations: 616

R-squared	0.305243	Mean dependent var	0.909789
Adjusted R-squared	0.298398	S.D. dependent var	0.665005
S.E. of regression	0.557019	Sum squared resid	188.9546
Durbin-Watson stat	1.110873		

(b) Consider OUTPUT 3 on page 6. Discuss the signs and significance of the estimated coefficients.

(4 marks)

(c) When the model in OUTPUT 3 was re-estimated by OLS, the following information was obtained:

Determinant residual covariance	5.669095

Using a formal statistical test, determine the model in OUTPUT 3 should be kept, or if the same model should be estimated by OLS. Be sure to carefully state your null and alternative hypotheses. Is your conclusion sensitive to your choice of significance level?

(5 marks)

OUTPUT 2

System: SYS01
Estimation Method: Iterative Seemingly Unrelated Regression
Date: 13/04/10 Time: 13:20
Sample: 1 616
Included observations: 616
Total system (balanced) observations 1232
Simultaneous weighting matrix & coefficient iteration
Convergence achieved after: 1 weight matrix, 2 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-3.708936	1.064465	-3.484320	0.0005
C(2)	0.438941	0.067204	6.531510	0.0000
C(3)	0.099272	0.015497	6.405995	0.0000
C(4)	0.673529	0.392213	1.717254	0.0862
C(5)	0.454913	0.412638	1.102451	0.2705
C(6)	1.032313	0.609924	1.692528	0.0908
C(7)	1.770192	0.396792	4.461255	0.0000
C(11)	-0.752465	0.135072	-5.570841	0.0000
C(12)	0.079526	0.008528	9.325716	0.0000
C(13)	0.010248	0.001966	5.211478	0.0000
C(14)	0.464771	0.049769	9.338638	0.0000
C(15)	0.125161	0.052360	2.390373	0.0170
C(16)	0.102897	0.077394	1.329519	0.1839
C(17)	0.232670	0.050350	4.621075	0.0000
Determinant residual covariance		5.423959		

Equation: HREARN=C(1)+C(2)*EDUC+C(3)*EXPER+C(4)*UNION+C(5)
*MARRIED+C(6)*WHITE+C(7)*MALE
Observations: 616

R-squared	0.184373	Mean dependent var	6.229286
Adjusted R-squared	0.176337	S.D. dependent var	4.836833
S.E. of regression	4.389710	Sum squared resid	11735.16
Durbin-Watson stat	1.950503		

Equation: HRBENS=C(11)+C(12)*EDUC+C(13)*EXPER+C(14)*UNION
+C(15)*MARRIED+C(16)*WHITE+C(17)*MALE
Observations: 616

R-squared	0.305243	Mean dependent var	0.909789
Adjusted R-squared	0.298398	S.D. dependent var	0.665005
S.E. of regression	0.557019	Sum squared resid	188.9546
Durbin-Watson stat	1.110873		

(d) The results in OUTPUT 4 relate to the model estimated in OUTPUT 3. Discuss and interpret these results.

(3 marks)

(e) I used the following two commands:

SCALAR EARN1=C(1)+C(2)*@MEAN(EDUC)+C(3)*@MEAN(EXPER)+C(4)+C(6)+C(7)
SCALAR EARN2=C(1)+C(2)*@MEAN(EDUC)+C(3)*@MEAN(EXPER)+C(4)+C(7)

The following values were obtained: EARN1 = 7.46713 ; EARN2 = 6.60425
Carefully interpret what these two values tell us.

(5 marks)

Total: 20 marks

OUTPUT 3

System: SYS01
Estimation Method: Iterative Seemingly Unrelated Regression
Date: 13/04/10 Time: 13:26
Sample: 1 616
Included observations: 616
Total system (balanced) observations 1232
Simultaneous weighting matrix & coefficient iteration
Convergence achieved after: 3 weight matrices, 4 total coef iterations

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-3.441245	1.054492	-3.263415	0.0011
C(2)	0.442475	0.067243	6.580220	0.0000
C(3)	0.103058	0.015106	6.822432	0.0000
C(4)	0.704491	0.391565	1.799167	0.0722
C(6)	0.862880	0.586627	1.470918	0.1416
C(7)	1.884984	0.385648	4.887834	0.0000
C(11)	-0.673589	0.123212	-5.466918	0.0000
C(12)	0.080810	0.008490	9.518670	0.0000
C(13)	0.010288	0.001965	5.237075	0.0000
C(14)	0.465030	0.049830	9.332321	0.0000
C(15)	0.114608	0.050381	2.274804	0.0231
C(17)	0.239958	0.050246	4.775646	0.0000
Determinant residual covariance		5.450788		

Equation: HREARN=C(1)+C(2)*EDUC+C(3)*EXPER+C(4)*UNION+C(6)
*WHITE+C(7)*MALE
Observations: 616

R-squared	0.182590	Mean dependent var	6.229286
Adjusted R-squared	0.175890	S.D. dependent var	4.836833
S.E. of regression	4.390901	Sum squared resid	11760.81
Durbin-Watson stat	1.938794		

Equation: HRBENS=C(11)+C(12)*EDUC+C(13)*EXPER+C(14)*UNION
+C(15)*MARRIED+C(17)*MALE
Observations: 616

R-squared	0.303146	Mean dependent var	0.909789
Adjusted R-squared	0.297434	S.D. dependent var	0.665005
S.E. of regression	0.557402	Sum squared resid	189.5249
Durbin-Watson stat	1.106334		

OUTPUT 4

Wald Test:
System: SYS01

Test Statistic	Value	df	Probability
Chi-square	0.394486	1	0.5300

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(4) - C(14)	0.239460	0.381257

Restrictions are linear in coefficients.

END OF EXAMINATION