Department of Economics **University of Victoria**

ECON 545: Econometric Analysis
Term Test, October 2014

Instructor:	David Giles
Instructions:	Answer **ALL QUESTIONS**, and put all answers in the booklet provided
Time Allowed:	120 minutes (Total marks = 120 – *i.e., one mark per minute*)
Number of Pages:	**FIVE**

PART A:
*Select the **most appropriate** answer in each case. Each question is worth **3 marks**.* (*No explanation is need for full marks, but it will be taken into account if given.*)

Question 1:
If we use Ordinary Least Squares to estimate the regression model, $y = X\beta + \varepsilon$, where ***all*** of the usual assumptions are satisfied, and the model includes an intercept then:
(a) The average value of the OLS residuals will be zero.
(b) The average of the OLS fitted values will equal the sample average of the y values.
(c) The regressors will be uncorrelated with the residuals in the sample.
(d) All of the above.

Question 2:
The coefficient of determination (R^2) for the usual least squares regression model:
(a) Always lies between zero and one in value.
(b) Cannot increase when we add one or more regressors to the model.
(c) Cannot be greater than the "adjusted" R^2.
(d) Can be negative if there is no intercept in the model.

Question 3:
For a linear multiple regression model, satisfying all of the usual assumptions, estimated by OLS:
(a) The error vector has a scalar covariance matrix, but the residual vector does not.
(b) The covariance matrix for the error vector is non-singular, but the covariance matrix for the residual vector is singular.
(c) Apart from a scale factor, the covariance matrices of the error vector and of the least squares residual vector are both idempotent.
(d) All of the above.

Question 4:
The "power" of any statistical test:
(a) Is just one minus the probability of a "Type I" error.
(b) Is equal to the significance level chosen for the test, when the null hypothesis is true.
(c) Always increases as the sample size grows, because more information is being used.
(d) Can never be smaller than the chosen significance level.

Question 5:
When we construct a confidence interval for one of the coefficients in a regression model:
(a) This interval will be shorter if the sample size is reduced, other things being equal.
(b) This interval will be wider if the confidence level is reduced, other things being equal.
(c) This interval will be wider if it constructed asymmetrically about b_i, than if it symmetric.
(d) None of the above.

Question 6:

The Gauss-Markov Theorem tells us that, under appropriate assumptions, the least squares estimator of β in the usual linear regression model:

(a) Is a linear estimator, and therefore is "best".

(b) Has the smallest bias among all possible linear estimators for this parameter vector.

(c) Is most efficient among all possible linear and unbiased estimators of this parameter.

(d) Is most efficient among all possible unbiased estimators that have a Normal sampling distribution.

Question 7:

The Instrumental Variables (I.V.) estimator for the coefficient vector in a linear regression model is designed to be:

(a) "Best linear unbiased" if the regressors are correlated with the errors.

(b) At least weakly consistent, if the regressors are correlated with the errors.

(c) Asymptotically efficient, relative the OLS estimator, whether the regressors are correlated with the errors or not.

(d) Mean square consistent when the instruments are highly correlated with the errors.

Total: 21 marks

PART B:

*State whether each of the following is TRUE or FALSE, and BRIEFLY explain your answer. Each question is worth **4 marks**. Of these, **3 marks are given for the explanation**.*

Question 8:

(a) If we apply a Hausman test of the hypothesis that the errors in a regression model are asymptotically uncorrelated with the regressors, we would use I.V. estimation if the p-value for the test is large enough (say, greater than 10% or 20%).

(b) Suppose that all of the usual assumptions for our regression model hold, except that the errors are heteroskedastic. Let b_i denote the usual OLS estimator of β_i, the i^{th} element of β. Then $[b_1\ b_2]$ is an unbiased and weakly consistent estimator of $[\beta_1 \beta_2]$.

(c) The Wald test is more useful than is the usual F-test that we use for testing exact linear restrictions on the regression coefficient vector, β.

(d) Suppose that one of the regressors in a k-regressor multiple linear OLS regression model is a variable that takes the value "1" for just the first three observations, but is zero for all other observations. Then, $e_1 = -(e_2 + e_3)$, where the e_i's are the OLS residuals.

(e) If we fit a linear regression model using Instrumental Variables estimation, with an equal number of regressors and instruments, the residuals will sum to zero as long as the instrument matrix includes a column of 'ones'.

(f) One connection between the variance of an estimator and the mean squared error of that estimator is that the mean squared error cannot be less than the variance.

Total: 24 marks

PART C: *Answer all 3 questions in this section.*

Question 9:

The following EViews results relate to a model that explains the net worth of a cross-section of U.S. individuals in 1989, measured in thousands of dollars. The regressors are:

EDUC = years of education
MARRIED = dummy variable (= 1 if married; = 0 if not)
PYEARS = number of years in a pension plan
AGE = age, in years
AFRAM = dummy variable (=1 if African American; = 0 if not)
PCTSTCK = % of pension plan held in stocks
FINC101 = dummy variable (= 1 if family income > $100,000 p.a.; = 0, otherwise)

Dependent Variable: NET_WORTH
Method: Least Squares
Date: 10/22/14 Time: 10:05
Sample: 1 226
Included observations: 205

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-679.1467	259.2987	-2.619167	0.0095
EDUC	17.47498	6.184717	2.825511	0.0052
MARRIED	99.17807			0.0095
PYEARS	-3.350712	1.675122	-2.000280	0.0468
AGE	10.39204	3.982739	2.609270	0.0098
AFRAM	-96.92262	51.90103	-1.867451	0.0633
PCTSTCK	-0.198055	0.398787	-0.496644	0.6200
FINC101	180.2197	70.46655	2.557521	0.0113

R-squared		Mean dependent var	210.8534
Adjusted R-squared	0.158546	S.D. dependent var	242.1072
S.E. of regression	222.0869	Akaike info criterion	13.68226
Sum squared resid	9716547.	Schwarz criterion	13.81194
Log likelihood	-1394.431	Hannan-Quinn criter.	13.73471
F-statistic	6.491077	Durbin-Watson stat	1.952117
Prob(F-statistic)	0.000001		

(a) *Briefly* discuss the signs and significance of the regressors. Are there any surprises? What do you conclude from the "F-statistic" in the above results?

(9 marks)

(b) Calculate the value for the "missing" R^2 , and explain what this value tells us.

(6 marks)

(c) Construct a 95% confidence interval for the coefficient of AGE, and carefully interpret its meaning.

(6 marks)

(d) Test the hypothesis that the coefficient of EDUC equals 10, using a 2-sided alternative hypothesis, and a 5% significance level.

(6 marks)

(e) Use the information below to compute the standard error for the estimated coefficient of the MARRIED regressor.

(3 marks)

Coefficient Covariance Matrix

C	EDUC	MARRIED	PYEARS	AGE	AFRAM	PCTSTCK	FINC101
67235.83	-613.4542	378.6997	91.61959	-967.7661	-2188.994	-23.21021	-515.3223
-613.4542	38.25072	-4.959195	-0.692818	1.899751	-0.258175	-0.066537	-67.34109
378.6997	-4.959195	1435.012	-7.638729	-21.52264	518.7295	-0.757774	-80.66601
91.61959	-0.692818	-7.638729	2.806033	-1.781839	-7.284819	-0.008129	1.544245
-967.7661	1.899751	-21.52264	-1.781839	15.86221	25.86974	0.279502	16.78817
-2188.994	-0.258175	518.7295	-7.284819	25.86974	2693.717	0.130063	293.9980
-23.21021	-0.066537	-0.757774	-0.008129	0.279502	0.130063	0.159031	3.124250
-515.3223	-67.34109	-80.66601	1.544245	16.78817	293.9980	3.124250	4965.534

Total: 30 marks

Question 10:

Consider the linear multiple regression model, $y = X\beta + \varepsilon$, where all of the usual assumptions are satisfied, ***except*** that the regressors are random and correlated with the errors (even asymptotically). That is, $p\lim\left(n^{-1}X'\varepsilon\right) = \gamma \neq 0$, and γ is finite. We have available a set of k instrumental variables which are form the columns of the (n x k) matrix Z. The X and Z matrices satisfy the following conditions:

(i) $p\lim\left(n^{-1}X'X\right) = Q_{XX}$; positive-definite & finite.

(ii) $p\lim\left(n^{-1}Z'Z\right) = Q_{ZZ}$; positive-definite & finite.

(iii) $p\lim\left(n^{-1}Z'X\right) = Q_{ZX}$; positive-definite & finite.

(iv) $p\lim\left(n^{-1}Z'\varepsilon\right) = 0$.

In this case the I.V. estimator for β is $b_{IV} = (Z'X)^{-1}Z'y$.

(a) Show that the I.V. residual vector is $e_{IV} = W\varepsilon$, where $W = [I - X(Z'X)^{-1}Z']$.

4 marks

(b) Show that the sum of the squares of these I.V. residuals can be written as:

$$\varepsilon'\varepsilon - \varepsilon'X(Z'X)^{-1}Z'\varepsilon - \varepsilon'Z(X'Z)^{-1}X'\varepsilon + \varepsilon'Z(X'Z)^{-1}X'X(Z'X)^{-1}Z'\varepsilon .$$

6 marks

(c) Using assumptions (i) – (iv) and Khintchine's Theorem, ***prove*** that $s_{IV}^2 = (e_{IV}'e_{IV})/n$ is a weakly consistent estimator of σ^2 (the variance of the error term in the model).

[**Hint:** the proof follows the same lines as the proof of the weak consistency of s^2 for the OLS case.]

14 marks

Total: 24 marks

Question 11:

(a) Suppose that we have the following regression model, which satisfies *all* of the usual assumptions:

$$y = X_1\beta_1 + X_2\beta_2 + \varepsilon \, , \tag{1}$$

and suppose that (from some other source) we have an unbiased estimator, say $\hat{\beta}_1$, for the β_1 vector. We can get an estimator of β_2 by regressing $(y - X_1\hat{\beta}_1)$ on X_2, using OLS. ***Prove*** that this will yield an *unbiased* estimator of β_2.

(7 marks)

(b) Suppose that we estimate the following regression model, which satisfies *all* of the usual assumptions, by OLS:

$$y = X\beta + \varepsilon \, . \tag{2}$$

Let the residual vector be denoted e.

We then estimate the following model by OLS:

$$y = X\beta + \alpha e + u \, . \tag{3}$$

Prove that the estimated value for the β vector from (3) will be the same as would be obtained in equation (2).

[**<u>Hint</u>:** If we apply OLS estimation to (1), then $b_1 = [X_1{}'M_2X_1]^{-1}X_1{}'M_2y$,

where $M_2 = I - X_2(X_2{}'X_2)^{-1}X_2{}'$.]

(6 marks)

(c) Using the result that $y = \hat{y} + e$, show that $e'y = e'e$.

(3 marks)

(d) Using the same approach as in part (b) above, ***prove*** that the value of the OLS estimator of α in equation (3) is one.

(5 marks)

Total: 21 marks

<u>END OF TEST</u>