Department of Economics **University of Manitoba**

ECON 7010: Econometrics I
Midterm, Oct. 31, 2012

Instructor: **Ryan Godwin**
Instructions: **Answer ALL QUESTIONS, and put all answers in the booklet provided**
Time Allowed: **75 minutes (Total marks = 75 – *i.e., one mark per minute*)**
Number of Pages: **5**

PART A:
Select the most appropriate answer in each case. Each question is worth 3 marks. (No explanation is needed to obtain 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 OLS residuals will sum to zero.
(b) There will be an equal number of positive and negative residuals.
(c) The "adjusted" R^2 will be positive.
(d) All of the above.

Question 2:
The Gauss-Markov Theorem tells us that, under appropriate assumptions about the model:

(a) The least squares estimator of β in the usual linear regression model has the smallest possible variance, and therefore is 'best'.
(b) The least squares estimator of β in the usual linear regression model has the smallest mean squared error among all possible linear estimators for this parameter vector.
(c) The least squares estimator of β in the usual linear regression model has the smallest mean squared error among all possible unbiased estimators of this parameter vector.
(d) The least squares estimator of β in the usual linear regression model is most efficient among all possible linear and unbiased estimators of this parameter vector.

Question 3:
The probability of a 'Type I' error for any statistical test:

(a) Is just one minus the power of the test.
(b) Is equal to one minus the probability of the 'Type II' error for the test.
(c) Will always be less that the probability of the 'Type II' error for the test.
(d) Is chosen in advance by the researcher, who then tries to design the test so as to make the test as powerful as possible.

Question 4:
When we construct a confidence interval for one of the coefficients in a regression model:

(a) This interval will generally be wider if the sample size is increased, other things equal.
(b) This interval will be wider if the confidence level is increased, other things being equal.
(c) The interval will be wider if the standard error for the estimated coefficient is smaller, other things being equal.
(d) All of the above.

Question 5:
Suppose we fit a linear regression model by Instrumental Variables estimation. The purpose of using this estimator is:

(a) To obtain an estimator that is mean square consistent, even though it may be biased in finite samples.
(b) To obtain an estimator that is weakly consistent and asymptotically efficient.
(c) To obtain an estimator that is at least weakly consistent.
(d) To obtain an estimator that is asymptotically unbiased and asymptotically efficient.

Question 6:
If we wish to test the validity of "J" independent linear restrictions on the coefficients of the usual linear regression model, and if ***all*** of the usual assumptions about that model are satisfied, then:

(a) We can use a Wald test, and this test will be valid even if the sample size is small.
(b) We can use a Wald test, and it will be uniformly most powerful.
(c) We can use an F-test, and if the restrictions are false the distribution of the test statistic will be F with J and $(n - k)$ degrees of freedom.
(d) We can use an F-test, and this test will be valid even if the sample size is small.

PART B:

*State whether each of the following is TRUE or FALSE, and BRIEFLY explain your answer. Each question is worth **5 marks**. Of these, **4 marks** are given for the explanation.*

Question 7: The OLS regression 'line' passes through the mean of the sample. This implies that there must be an equal number of positive residuals and negative residuals if the sample size, n, is an even number.

Question 8: 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%).

Question 9: The I.V. estimator is always more efficient than the OLS estimator.

Question 10: Suppose we are using a Wald test to determine whether the restriction $R\beta = q$ is true. If the restriction is true, then $Rb - q$ will be large compared to the variance of $Rb - q$ (where b is the OLS estimator for β). The Wald statistic will then be "extreme" compared to the Wald distribution, and we will accept the null hypothesis.

PART C: *Answer all questions.*

Question 11. (15 marks total)

Consider two models:

Model 1: $y = \beta_1 X_1 + \varepsilon,$

Model 2: $y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon.$

Assume that:

(i) $plim\left(\frac{1}{n} X_1' X_1\right) = Q_{11}$; positive-definite & finite.

(ii) $plim\left(\frac{1}{n} X_2' X_1\right) = Q_{21}$; positive-definite & finite.

(iii) $plim\left(\frac{1}{n} X_2' X_2\right) = Q_{22}$; positive-definite & finite.

(iv) $plim\left(\frac{1}{n} Z' \varepsilon\right) = 0.$

(a) Suppose that Model 1 is the true model, but that you fit Model 2. Show that the OLS estimators obtained from fitting Model 2 are unbiased and consistent.

(5 marks)

(b) Now suppose that Model 2 is the true model, but that you fit Model 2. Show that the OLS estimators obtained from fitting Model 1 are biased and inconsistent.

(5 marks)

(c) Argue that R^2 cannot decrease when a regressor is added to the model.

(3 marks)

(d) Briefly discuss the merit of a "general-to-specific" strategy for model selection.

(2 marks)

Question 12. (12 marks total)

Consider the standard regression model, $y = X\beta + \varepsilon$, where all of the usual assumptions (including normality of the errors) are satisfied. Consider the following "family" of possible estimators for the variance of the error term: $\tilde{\sigma}^2 = (e'e)/(n - c)$, where "$e$" is the OLS residual vector, n is the sample size, and "c" is some scalar constant.

(a) What is the bias of $\tilde{\sigma}^2$? What determines the sign (direction) of the bias?

(4 marks)

(b) What is the variance of $\tilde{\sigma}^2$?

(4 marks)

Hint: Recall that $\frac{(n-k)s^2}{\sigma^2}$ is distributed as χ^2 with $(n - k)$ degrees of freedom; $E\left[\chi^2_{(v)}\right] = v$; and $var.\left[\chi^2_{(v)}\right] = 2v$.

(c) Prove that this family of estimators are weakly consistent.

(4 marks)

Question 13. (10 marks total)

The following EViews results are for a model explaining the demand for gasoline in the U.S. The variables are: POP = Population; G = Sales of gasoline (real $); Y = Real *per capita* disposable income; PG = Gasoline price index; PN = Non-durables price index; PD = Durables price index; PNC = New automobiles price index; PUC = Used automobiles price index.

Dependent Variable: G/POP

Method: Least Squares

Sample: 1960 1995

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.228689	0.180710	6.799225	0.0000
Y	9.881010	0.903674		0.0000
PG/PN	-0.082130	0.012507	-6.566701	0.0000
PNC/PD		0.070176	-4.646567	0.0001
PUC/PD	-0.079419	0.023114	-3.436051	

R-squared	0.958817	Mean dependent var	1.006903
Adjusted R-squared	0.953504	S.D. dependent var	0.140776
S.E. of regression	0.030356	Akaike info criterion	-4.023430
Sum squared resid		Schwarz criterion	-3.803497
Log likelihood	77.42174	F-statistic	180.4366

(a) Calculate the estimated coefficient for (PNC / PD) and the t-statistic for Y.

(2 marks)

(b) Calculate the value for the sum of the squared residuals.

(2 marks)

(c) Construct a 99% confidence interval for the coefficient of the relative price of gasoline. (A t-distribution with 30 degrees-of-freedom contains a tail probability of 0.05% when $t_C = 2.75$). How many degrees-of-freedom do the above t-statistics have?

(3 marks)

(d) In the bottom right of the table, we see the "F-statistic" is 180.4366. What is the null and alternative hypothesis associated with this F-test? What are you likely to conclude from the value of this statistic?

(3 marks)