Introduction to Econometrics (Econ 107)
Drake University, Spring 2010
William M. Boal

Signature: ______________________

Printed name: ______________________

MIDTERM EXAMINATION #4 VERSION A
"Multiple Regression With Time-Series Data"
April 23, 2010

INSTRUCTIONS: This exam is closed-book, closed-notes. You may use a calculator for this exam, but not a graphing calculator or a calculator with alphabetical keys. Point values for each question are noted in brackets. Tables of the t distribution, F distribution, and chi-square distribution are attached.

NOTATION: In this exam, $\hat{\beta}_j$ denotes the least-squares coefficient estimators of the linear model $y_t = \beta_1 + \beta_2 x_{t2}, + \ldots + \beta_K x_{tK} + \varepsilon_t$. The least-squares fitted value is denoted $\hat{y}_t$. The least-squares residual is denoted $\hat{\varepsilon}_t$. The array denoted X includes all regressors in all time periods in the sample. The sample size is denoted T. The true or population value of the variance of the unobserved error term ε_t is denoted σ^2. The (unbiased) least-squares estimator of σ^2 is denoted $\hat{\sigma}^2$. The sample mean of y is denoted $\bar{y}$.

I. MULTIPLE CHOICE: Circle the one best answer to each question. Use margins for scratch work [2 pts each—14 pts total]

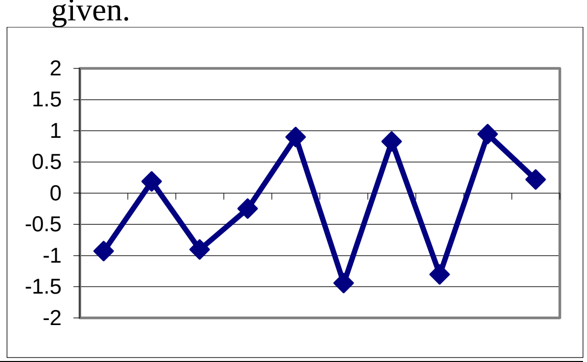
(1) Which is a static model?
a. $HS_t = 6.4 - 0.2\ R_{t-1} + \varepsilon_t$.
b. $HS_t = 2.3 - 1.1\ UNEMP_{t-1} - 0.7\ UNEMP_{t-1} + \varepsilon_t$.
c. $HS_t = 3.1 - 1.7\ UNEMP_t - 0.3\ R_t + \varepsilon_t$.
d. $HS_t = 2.5 - 2.7\ UNEMP_{t-1} + \varepsilon_t$.

(2) According to the model

$$y_t = 5.9 + 0.08\ t + 7.1\ x_t ,$$

if x remains constant, then y increases from one period to the next, on average, by
a. 0.08 percent.
b. 8 percent .
c. 8 units .
d. 0.08 units .
e. Cannot be determined from information given.

(3) The time series ε_t graphed below has mean zero. It appears to be
a. positively serially-correlated.
b. negatively serially-correlated.
c. serially uncorrelated.
d. Cannot be determined from information given.



(4) If there is *negative* serial correlation, the Durbin-Watson test statistic is most likely

a. less than zero.
b. approximately 0.
c. between 0 and 1.
d. approximately 1.
e. between 0 and 2.
f. approximately 2.
g. between 2 and 4.
h. approximately 4.

(5) If the random process u_t has a unit root, then

a. $u_t = 1$.
b. u_t wanders away from its mean.
c. $E(u_t) = 1$.
d. $Var(u_t) = 1$.
e. the square root of $u_t = 1$.
f. All of the above.

(6) Suppose y_t and x_t are two independent random walks. If the equation

$$y_t = \beta_1 + \beta_2 x_t$$

is estimated by ordinary least squares, it will yield

a. an excessively small (in absolute value) t-statistic for β_2 .
b. an excessively large (in absolute value) t-statistic for β_2 .
c. a valid t statistic for β_2 .
d. an R-square value close to zero.

(7) Suppose we estimate the following vector autoregression (VAR) model:

$$GDP_t = \alpha_1 + \alpha_2 GDP_{t-1} + \alpha_3 M2_{t-1} + \varepsilon_t$$
$$M2_t = \beta_1 + \beta_2 M2_{t-1} + \beta_3 GDP_{t-1} + \upsilon_t$$

To test the hypothesis that GDP Granger-causes M2, we must use the t-statistic for

a. α_1 .
b. α_2 .
c. α_3 .
d. β_1 .
e. β_2 .
f. β_3 .

II. SHORT ANSWER: Please write your answers in the boxes on this question sheet. Use margins for scratch work.

(1) [Stochastic processes: 3 pts] Let ε_t denote an independent identically-distributed (IID) process, and consider the stochastic processes defined by the equations below.

(i) $u_t = 0.4\, u_{t-1} + \varepsilon_t$.
(ii) $u_t = \varepsilon_t + 0.6\, \varepsilon_{t-1} - 0.1\, \varepsilon_{t-2}$.
(iii) $u_t = 6.7 + u_{t-1} + \varepsilon_t$.
(iv) $u_t = \varepsilon_t + 0.5\, \varepsilon_{t-1}$.
(v) $u_t = u_{t-1} + \varepsilon_t$.
(vi) $u_t = 0.5\, u_{t-1} + 0.3\, u_{t-2} + \varepsilon_t$.

Match the name of the stochastic process with the equation by inserting a roman numeral in each box below.

a. Random walk with drift.	
b. First-order moving-average process or MA(1).	
c. Second-order autoregressive process or AR(2).	

(2) [Assumptions: 7 pts] Consider the following possible characterizations of the error term ε_t, in mathematical notation.

(0) $\varepsilon_t \sim N(0, \sigma^2)$.
(i) $E(\varepsilon_t) = 0$, for all t.
(ii) $Var(\varepsilon_t) = 0$, for all t.
(iii) $Cov(\varepsilon_t, \varepsilon_s) = 0$, for all $s \neq t$.
(iv) $Var(\varepsilon_t) = \sigma^2$, for all t.
(v) $E(\varepsilon_t \mid x_t) = 0$, for all t.
(vi) $E(\varepsilon_t \mid X) = 0$, for all t.
(vii) $Cov(\varepsilon_t, \varepsilon_{t-h}) \to 0$ as $h \to \infty$, for all t.
(viii) $E(\varepsilon_t) = E(\varepsilon_s)$, $Var(\varepsilon_t) = Var(\varepsilon_s)$, and $Cov(\varepsilon_t, \varepsilon_{t-h}) = Cov(\varepsilon_s, \varepsilon_{s-h})$ for all s and t.

Match the mathematical characterization to the verbal characterization by inserting a roman numeral in each box below.

a. The error term is homoskedastic.	
b. The error term is not serially-correlated.	
c. The regressors are strictly exogenous.	
d. The regressors are contemporaneously exogenous.	
e. The error term is asymptotically uncorrelated.	
f. The error term is covariance-stationary.	
g. The error term is normally-distributed with mean zero.	

(3) [Properties of LS: 5 pts] For the model $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$, ordinary least squares yields consistent estimators of β_1 and β_2 if (write *yes* or *no*)

a. ε_t is serially-correlated, but stationary and weakly dependent, and x_t is weakly exogenous.	
b. ε_t is an independent, identically-distributed process, and x_t is strictly exogenous.	
c. x_t and y_t are integrated processes but are not co-integrated.	
d. x_t and y_t are co-integrated processes.	
e. ε_t is a random walk.	

III. PROBLEMS: Please write your answers in the boxes on this question sheet. Use margins for scratch work.

(1) [Finite distributed lag: 6 pts] Suppose we have estimated the time-series model

$$quantity_t = 3.5 + 2.1\ price_t + 1.2\ price_{t-1} + 0.6\ price_{t-2} + \varepsilon_t\ .$$

a. Is this a *static* model or a *dynamic* model?

b. Compute the impact propensity (also called the "impact multiplier" or "short-run effect").

c. Compute the long-run propensity (also called the "long-run multiplier" or "long-run effect").

(2) [Breusch-Godfrey test: 12 pts] Suppose we have estimated the regression model

$$y_t = \beta_1 + \beta_2\, x_t + \beta_3\, y_{t-1} + \varepsilon_t$$

using 81 observations, but we fear that ε_t might have first-order serial correlation. To perform a Breusch-Godfrey test, we estimate an auxiliary regression.

a. What should be the dependent variable of the auxiliary regression?

b. What should be the regressors of the auxiliary regression?

(Note that this auxiliary regression must be estimated on observations 2 through 81 of the original data.) The R^2 value from this auxiliary regression is 0.055 . Test the null hypothesis of no serial correlation at 5% significance. Give the value of the test statistic, its degrees of freedom, the critical point from the appropriate table, and your conclusion (whether you can reject the null hypothesis).

c. Degrees of freedom =

d. Value of test statistic =

e. Critical point =

f. Can you reject the null hypothesis?

(3) [Quasi-differencing: 12 pts] Suppose we wish to estimate the time-series regression model

$$y_t = \beta_1 + \beta_2 x_t + u_t .$$

However, we believe the error term is serially correlated, following the process

$$u_t = 0.25 u_{t-1} + \varepsilon_t ,$$

where ε_t is an independent, identically-distributed random error term. We need to transform the data to eliminate the serial correlation. The table below shows the first three observations on y_t and x_t. Compute transformed values of the second and third observations using the Cochrane-Orcutt method.

Obs.	Raw data		Transformed data		
t	y_t	x_t	y_t	Replacement for intercept	x_t
1	24	8			
2	28	12			
3	35	16			

(4) [Random walk: 12 pts] Consider the following process:

$$u_t = 2.1 + u_{t-1} + \varepsilon_t .$$

where ε_t denotes an independent, identically-distributed series with $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = 0.53$. Assume $u_0 = 0$.

a. Find a formula in terms of t for the unconditional mean $E(u_t)$.

b. Find a formula in terms of t for the unconditional variance $Var(u_t)$.

c. Is u_t a *stationary* process or a *nonstationary* process?

d. Let $\Delta u_t = u_t - u_{t-1}$. Compute $E(\Delta u_t)$.

e. Compute $Var(\Delta u_t)$.

f. Suppose $u_{50} = 128$. Compute the forecast of u_{51}.

(5) Dickey-Fuller test: 6 pts] We wish to test whether the inflation has a unit root using an augmented Dickey-Fuller test. We have estimated the following by ordinary least squares, using a large sample, standard errors in parentheses.

$$\underset{}{\Delta \text{inflation}_t} = \underset{(0.31)}{0.93} - \underset{(0.10)}{0.16}\ \text{inflation}_{t-1} + \underset{(0.06)}{0.27}\ \Delta\text{inflation}_{t-1}$$

a. Which is the null hypothesis of the Dickey-Fuller test: that inflation *does not have* a unit root, or that inflation *has* a unit root?

b. Compute the Dickey-Fuller test statistic.

c. The critical point at 5% significance for the augmented Dickey-Fuller test statistic is –2.86. Can you reject the null hypothesis at 5% significance?

(6) [Forecasting, trends and seasonality: 6 pts] Suppose we have estimated the following model for sales, using 80 quarterly observations from the first quarter of 1990 to the fourth quarter of 2009.

$$\text{sales}_t = 88.3 + 0.2\ \text{trend} - 0.6\ q1_t - 0.2\ q2_t - 0.5\ q3_t + \varepsilon_t\ .$$

The regressor “trend” equals 1 in the first quarter of 1990, equals 2 in the second quarter of 1990, and so forth, and equals 80 in the fourth quarter of 2009. The regressors “q1,” “q2,” and “q3,” are quarterly dummy variables for the first, second and third quarters respectively. The error term ε_t is an independent, identically-distributed process with $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = \sigma^2$, constant.

a. If a dummy variable “q4” for the fourth quarter were also included, then what econometric problem would result?

b. Compute the forecast of sales in the first quarter of 2010.

c. Compute the forecast of sales in the second quarter of 2010.

(7) [Forecasting, forecast interval: 11 pts] We wish to use the equation

$$employ_t = \beta_1 + \beta_2 \, gdp_{t-1} + \beta_3 \, employ_{t-1} + \varepsilon_t$$

to forecast $employ_{181}$. The last observations in our sample are $gdp_{180} = 14.1$ and $employ_{180} = 218$. To compute the forecast, we first transform the data.

a. [2 pts] Should the dependent variable be transformed? If so, how?

b. [2 pts] Should the regressors be transformed? If so, how?

Suppose the equation has been estimated on the ***transformed data*** with the following results, standard errors in parentheses. The estimated variance of the error term is $\hat{\sigma}^2 = 13.75$.

$employ_t$	=	221.0	+	22.3	gdp_{t-1}	+	0.21	$employ_{t-1}$
		(1.5)		(1.2)			(0.04)	

c. [2 pts] Compute the forecast of $employ_{181}$.

d. [2 pts] Compute the standard error of forecast error.

e. [3 pts] Compute a 95% forecast interval for $employ_{181}$. Assume $\varepsilon_t \sim N(0, \sigma^2)$.

(8) [Forecasting, AR model: 6 pts] Suppose we have estimated the following model.

$$unemp_t = 3.3 + 0.6 \, unemp_{t-1} - 0.2 \, unemp_{t-2} + \varepsilon_t \, .$$

where ε_t denotes an independent, identically-distributed process with $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = \sigma^2$, constant. In our data set, $unemp_{T-1} = 5$ and $unemp_T = 4.5$. Compute the following forecast values.

a. Compute the forecast of $unemp_{T+1}$.

b. Compute the forecast of $unemp_{T+2}$.

c. Compute the limit of the forecast $unemp_{T+h}$ as h approaches infinity. [Hint: this is the unconditional mean of $unemp_t$.]

[end of exam]