QUANTITATIVE ECONOMICS - FALL 2016
Answer any FIVE questions.

1. (a) Consider the regression model

$$Y_i = \beta X_i + u_i \quad i = 1, \ldots, n$$

where u_i and X_i satisfy the standard assumptions. Let $\overline{\beta} = \frac{\overline{Y}}{\overline{X}}$, where $\overline{Y}$ and $\overline{X}$ are the sample means of Y_i and X_i, respectively.

i. Show that $\overline{\beta}$ is a linear function of $Y_1, Y_2, ..., Y_n$.

ii. Show that $\overline{\beta}$ is conditionally unbiased.

iii. Derive the variance of $\overline{\beta}$.

(b) Consider the regression model $Y_i = \beta X_i + \varepsilon_i$. Suppose that a researcher mistakenly applies the least squares method to $X_i = \delta Y_i + u_i$ and then uses $\frac{1}{\hat{\delta}}$ as an estimator for β. Find the bias of $\frac{1}{\hat{\delta}}$.

2. Consider the regression model

$$y_i = \alpha + \beta x_i + u_i, \quad i = 1, 2, \ldots, n$$

(a) Suppose we assume that $\alpha = 0$ and hence $y_i = \beta x_i + u_i$

i. Write down the moment condition(s) used to derive the method of moments estimator for β.

ii. Write down the least-squares objective function to derive the OLS estimator for β.

iii. Use the equation from ii. to derive the estimator for β.

(b) Suppose we assume that $\beta = 0$ and hence $y_i = \alpha + u_i$

i. Write down the moment condition(s) used to derive the method of moments estimator for α.

ii. Write down the least-squares objective function to derive the OLS estimator for α.

iii. Use the equation from ii. to derive the estimator for α.

3. Suppose you collect data from a survey on wages, education, experience, and gender. In addition, you ask for information about marijuana usage. The original question is: "On how many separate occasions last month did you smoke marijuana?"

(a) Write an equation that would allow you to estimate the effects of marijuana usage on wage, while controlling for other factors. You should be able to make statements such as, "Smoking marijuana five more times per month is estimated to change wage by x%."

(b) Write a model that would allow you to test whether drug usage has different effects on wages for men and women. How would you test that there are no differences in the effects of drug usage for men and women?

(c) Suppose you think it is better to measure marijuana usage by putting people into one of four categories: nonuser, light user (1 to 5 times per month), moderate user (6 to 10 times per month), and heavy user (more than 10 times per month). Now, write a model that allows you to estimate the effects of marijuana usage on wage.

(d) Using the model in part (c), explain in detail how to test the null hypothesis that marijuana usage has no effect on wage. Be very specific and include a careful listing of degrees of freedom.

4. Consider the following data generating process

$$Y_t = c + \phi_2 Y_{t-2} + \varepsilon_t$$

where ε_t is a white noise process. Assuming $0 < \phi_2 < 1$, answer the following:

(a) Derive the expected value of this process.

(b) Derive the variance of this process.

(c) Derive the covariance of this process for $j = 1, 2, 3$.

(d) Derive the autocorrelation function of this process for $j = 1, 2, 3$.

5. In a fixed effects model, we can consider several ways to control for the individual effects which may be correlated with the regressors. Three cases being the within estimator, the least-squares dummy variable estimator (LSDV) and the first difference estimator. Consider the three methods, show that

   (a) The within estimator and the LSDV estimator are equivalent.

   (b) When $T = 2$, the within estimator and the first difference estimator are equivalent.

6. Suppose we have data for U.S. interest rate data over the period 1960:Q1 to 2008:Q1 and we are interested in estimating a quarterly model of spread between a long-term and a short-term interest rate. Specifically, the interest rate spread (s) can be formed as the difference between the interest rate on a 10-year U.S. government bonds (r10) and the rate on a three-month treasury bills (Tbill) as

$$s_t = r10_t - Tbill_t$$

   Suppose the interest rate spread in the fourth quarter of 2007 is 0.75, the interest rate spread in the first quarter of 2008 is 1.51 and the interest rate spread in the second quarter of 2008 is 1.61. Use the EViews output to answer the following questions:

Dependent Variable: S
Method: Least Squares
Date: 12/30/09 Time: 10:28
Sample (adjusted): 1960Q3 2008Q1
Included observations: 191 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.387620	0.285227	4.864963	0.0000
AR(1)	1.109152	0.070829	15.65953	0.0000
AR(2)	-0.245386	0.070792	-3.466283	0.0007

R-squared	0.805374	Mean dependent var	1.378220
Adjusted R-squared	0.803304	S.D. dependent var	1.210806
S.E. of regression	0.536998	Akaike info criterion	1.609936
Sum squared resid	54.21291	Schwarz criterion	1.661019
Log likelihood	-150.7489	F-statistic	388.9784
Durbin-Watson stat	1.960866	Prob(F-statistic)	0.000000
Inverted AR Roots	.80	.31	

   (a) Write the estimated equation given in the Eviews output.

   (b) Obtain the one-step ahead forecast.

   (c) Obtain the one-step ahead forecast error.

   (d) Obtain the two-step ahead forecast.

   (e) Obtain the three-step ahead forecast.

7. A variable Q is determined by the model

$$Q = \beta_1 + \beta_2 X + v$$

   where X is a variable and v is a disturbance term that satisfies the regression model assumptions. The dependent variable is subject to measurement error and is measured as Y where

$$Y = Q + r$$

   and r is the measurement error, distributed independently of v. Describe analytically the consequences of using OLS to fit this model if:

(a) The expected value of r is not equal to zero (but r is distrubted independently of Q).

(b) r is not distributed independently of Q (but its expected value is zero).

8. Suppose I am thinking about figuring out how much my rent will be given my income. I have gathered data on Washington, DC apartment rents ($/Year) and the incomes of the people who live in those apartments ($/Year), as well as the square footage of the apartment and whether it has an unobstructed view (Dummy Variable). Two regressions have been run and the results are reported in the table below.

Regressor/Dependent	Rent	$\widehat{u}^2$
	Model 1	Model 2
Intercept	5455.48	-361594.30
	(602.77)	(5330126.00)
Income	0.0635	305.6033
	(0.0143)	(127.2458)
Income^2	-0.0002	40.3662
	(0.0084)	(19.3689)
SquareFoot	0.0234	-59633.25
	(0.0074)	(28966.258)
View	1000.26	-26988.3698
	(206.93)	(30599.358)
Observations	108	108
R^2	0.1555	0.0516

(a) Interpret the impact of income on yearly rent?

(b) Conduct a test of no regression in model (1) at the 5% level. Be sure to write the null and alternative hypothesis correctly.

(c) Explain why the error variance of the model run in model (1) may be heteroscedastic with regard to income.

(d) Consider model (2). Here the dependent variable is the square of the residuals from model (1). What type of test can be implemented from model (2)? Perform this test at the 10% level.

(e) Does the test in (e) specifically test if the error variances are heteroscedastic with regards to income? Why or why not?