QUANTITATIVE ECONOMICS - FALL 2018
Answer any FIVE questions.

1. 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.

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

   (b) Show that $\overline{\beta}$ is conditionally unbiased.

   (c) Derive the variance of $\overline{\beta}$.

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. You are given the regression model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

   You want to test the null hypothesis

$$\begin{aligned} H_0 &: \beta_1^3 \beta_2 = \beta_3^2 \\ H_A &: \beta_1^3 \beta_2 \neq \beta_3^2 \end{aligned}$$

Show how you would construct the test statistic.

8. Consider the following regression with respect to female labor participation (y) on years of education (x):

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

where y is binary (0 or 1, 1 signifying the woman chooses to work) and x is measured in years (≥ 0). Assume that after OLS, we obtain the parameters of the model as

$$\begin{array}{rcc} y_i & = & -0.146 + 0.038 x_i + u_i \\ & & (0.121) \quad (0.014) \end{array}$$

where the number below the estimated coefficient in parantheses is the standard error of the estimate. With the above information, answer the following:

(a) Interpret the estimated coefficient for α. Does this seem reasonable?

(b) Interpret the estimated coefficient for β. Does this seem reasonable?

(c) Test the null that the coefficient in part (a) is zero. Test the null that the coefficient in part (b) is zero.

(d) Draw a figure with probability of labor force participation on the vertical axis and years of education on the horizontal axis. Draw and label the regression line.

(e) How many years of education does it take before labor force participation is possible (positive probability)? How many years of education does it take for labor force participation to be imminent?