QUANTITATIVE ECONOMICS - SPRING 2017
Answer any FIVE questions.

1. Consider the regression $Y = X\beta + u$ where X is $n \times k$ matrix and β is a $k \times 1$ vector. $E(u_i) = 0$; $E(u_i^2) = \sigma^2$ and $E(u_i u_j) = 0$ if $i \neq j$. Let $\widehat{\beta}$ denotes the least squares estimator of β.
   (a) Show that $\widehat{\beta}$ is conditionally unbiased.
   (b) Derive the conditional variance of $\widehat{\beta}$.
   (c) Show that $\widehat{\beta}$ is a consistent estimator. (Hint: Show that $Var(\widehat{\beta}) \to 0$ as $n \to \infty$)

2. Consider the following data generating process

$$Y_t = c + \phi_3 Y_{t-3} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

   where ε_t is a white noise process. Assuming $0 < \phi_3 < 1$ and $0 < \theta_1 < 1$, 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 $j = 1, 2, 3, 4, 5$.
   (d) Derive the autocorrelation function of this process for $j = 1, 2, 3, 4, 5$.

3. 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.

4. 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 0.95, the interest rate spread in the first quarter of 2008 is 2.21 and the interest rate spread in the second quarter of 2008 is 1.11. 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.

5. Consider the Eviews output listed below on the regression of log test scores on class size.

Dependent Variable: LNTESTSCORES
Method: Least Squares
Date: 08/14/08 Time: 08:05
Sample: 1 3733
Included observations: 3733

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.890871	0.010629	366.0531	0.0000
CLASSSIZE	0.002198	0.000436	5.043496	0.0000

R-squared	0.006772	Mean dependent var	3.942173
Adjusted R-squared	0.006505	S.D. dependent var	0.189056
S.E. of regression	0.188440	Akaike info criterion	-0.499534
Sum squared resid	132.4870	Schwarz criterion	-0.496199
Log likelihood	934.3802	F-statistic	25.43685
Durbin-Watson stat	1.716947	Prob(F-statistic)	0.000000

(a) Interpret the coefficient on class size.

(b) Show the formula used to calculate Residual Sum of Squares (SSR). Give the value of SSR.

(c) Show the formula used to calculate Total Sum of Squares (SST). Give the value of SST.

(d) Show the formula used to calculate Explained Sum of Squares (SSE). Give the value of SSE.

(e) Suppose we multiplied the log of test scores by 10 (for all observations). What will happen to SSR, SSE, SST and R^2?

(f) Suppose we multiplied the class size variable by 10 (for all observations). What will happen to SSR, SSE, SST and R^2?

6. After reading the textbook's analysis of test scores and class size, an educator comments: "In my experience, student performance depends on class size, but not in the way your regressions say. Rather, students do well when class size is less than 20 students and do very poorly when class size is greater than 25. There are no gains from reducing class size below 20 students, the marginal benefit is constant in the region between 20 and 25 students, and there is no loss to increasing class size when it is already greater than 25". The educator is describing a "threshold effect" in which performance is constant for class sizes less than 20, increases constantly as class size increases for class sizes between 20 and 25, and then is constant again for class sizes greater than 25. To model these threshold effects, define the binary variables

$$\begin{aligned} STRs &= 1 \text{ if } STR < 20, \text{ and } STRs = 0 \text{ otherwise;} \\ STRm &= 1 \text{ if } 20 \leq STR \leq 25, \text{and } STRm = 0 \text{ otherwise;} \\ STRl &= 1 \text{ if } STR > 25, \text{ and } STRl = 0 \text{ otherwise.} \end{aligned}$$

(a) Use the binary variables and STR to formulate a regression model (a single equation) that allows for each class-size group to have a different regression line.

(b) Based on your regression model in (a), formulate an appropriate test that all class-size groups have the same regression line against the alternative that at least one class-size group has a different regression line to the other two.

(c) Use the binary variables and STR to formulate test score as a continuous function of the student-teacher ratio that assumes constant but possibly different marginal effects of the student-teacher ratio on test score across these three different class-size groups

(d) Based on your regression model in (c), formulate an appropriate test of the educator's claim of at least one threshold effect against an alternative that there are no threshold effects.

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

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

Show how you could construct the test statistic.

8. Consider the linear model

$$y_i = x_i \beta + u_i$$

where $E(u|x) \neq 0$ but there exists a vector of variables z such that $E(u|z) = 0$. Assuming that we have more number of instruments than the endogenous variables ($L > K$), and using the moment condition $E(zu) = 0$, derive the explicit solution for the generalized method of moments estimator $\widehat{\beta}_{GMM}$.