**QUANTITATIVE ECONOMICS - SPRING 2016**
**Answer any FIVE questions.**

1. (a) Prove the result that the $R^2$ associated with a restricted least squares estimator is never larger than that associated with the unrestricted estimator.

   (b) 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

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

   Show how you would construct the test statistic.

2. Suppose you have the nonlinear regression

$$y_i = \alpha(1 + x_i^{\beta}) + u_i, \quad E[u_i|x_i] = 0$$

   and IID data $\{(x_i, y_i)\}_{i=1}^{n}$. Describe in detail how you could estimate $\alpha$ and $\beta$?

3. (a) Consider the regression $Y = X\beta + u$ where $X$ is $n \times k$ matrix and $\beta$ 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 $\beta$.Derive the conditional variance of $\widehat{\beta}$.

   (b) In the least squares regression of $\mathbf{y}$ on a constant and $\mathbf{X}$, to compute the regression coefficients on $\mathbf{X}$, we can first transform $\mathbf{y}$ to deviations from the mean $\overline{y}$ and, likewise, transform each column of $\mathbf{X}$ to deviations from the respective column mean; second, regress the transformed $\mathbf{y}$ on the transformed $\mathbf{X}$ without a constant. Do we get the same result if we only transform $\mathbf{y}$? What if we only transform $\mathbf{X}$?

4. Consider the following data generating process

$$Y_t = \varepsilon_t + \theta_3 \varepsilon_{t-3}$$

   where $\varepsilon_t$ is a white noise process. Assuming $-1 < \theta_3 < 0$, 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 $j = 1, 2, \ldots, J$. Note that $j$ represents the number of periods between $Y_t$ and $Y_{t-j}$.

   (d) Derive the autocorrelation function for all $j = 1, 2, \ldots, J$.

   (e) Plot the autocorrelation function with the information you derived above.

5. Consider the following data generating process

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3}$$

   where $\varepsilon_t$ is a white noise process and $|\phi_1 + \phi_2| < 1$. Without taking a first difference, answer the following:

   (a) Find the h-step-ahead forecast for $h = 1, 2, ..., k$.

   (b) Find the h-step-ahead forecast error for $h = 1, 2, ..., k$.

   (c) Find the h-step-ahead forecast error variance for $h = 1, 2, ..., k$.

6. 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.

7. Suppose we are interested in the relationship of the union status variable $Y$ (= 1; if in union, = 0, if not in union) to the conditioning variables $X_1$ = gender (1 if female, 0 if male), and $X_2$ = marital status (1 if married, 0 if not). Table below gives the coefficient estimates obtained in **(i)** Least squares regression of $Y$ on $X_1$ and $X_2$, and, **(ii)** Nonlinear least-squares estimates of the logistic regression model $E(Y|X_1, X_2) = G(Z)$, where $Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2$, and $G(Z) = e^Z/(1 + e^Z)$. In parenthesis are the conventional standard errors of the coefficient estimates. Also tabulated are the means of the conditioning variables (regressors).

| | Linear regression | Logistic model | Sample means |
|---|---|---|---|
| 1 | 0.192<br>(0.03) | −1.65<br>(0.26) | 1.00 |
| $X_1$ | −0.13<br>(0.03) | −1.01<br>(0.28) | 0.46 |
| $X_2$ | 0.08<br>(0.03) | −0.74<br>(0.29) | 0.66 |

   (a) Consider an single man and a married woman. According to the estimated logistic function, which of them has the higher probability of being a union member? How much larger?

   (b) Determine whether the linear function gives approximately the same answer to those questions above.

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 ($\geq 0$). Assume that after OLS, we obtain the parameters of the model as

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

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 $\alpha$. Does this seem reasonable?

   (b) Interpret the estimated coefficient for $\beta$. 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?