QUANTITATIVE ECONOMICS - FALL 2019

Answer any FIVE questions.

1. Consider the single equation model
$$y_i = z_i\beta + e_i$$
where y_i and z_i are both real-valued (1×1) and $E(e_i|z_i) \neq 0$. Let $\widehat{\beta}$ denote the IV estimator of β using as an instrument a dummy variable d_i (takes only the values 0 and 1). Derive the explicit expression for the IV estimator.

2. Consider the following data generating process
$$Y_t = \varepsilon_t + \theta_2\varepsilon_{t-2}$$
where ε_t is a white noise process with $V(\varepsilon_t) = \sigma^2$. Assuming $-1 < \theta_2 < 0$, 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 all j. Note that j represents the number of periods between Y_t and Y_{t-j}.
   (d) Derive the autocorrelation function for all j.
   (e) Plot the autocorrelation function with the information you derived above.

3. Consider a product market with a supply function $Q_i^s = \beta_0 + \beta_1 P_i + u_i^s$, a demand function $Q_i^d = \gamma_0 + u_i^d$, and a market equilibrium condition $Q_i^s = Q_i^d$, where u_i^s and u_i^d are mutually independent *i.i.d.* random variables, both with a mean of zero and variance σ_s^2 and σ_d^2, respectively.
   (a) Show that P_i and u_i^s are correlated.
   (b) Show that the OLS estimator of β_1 is inconsistent.
   (c) How would you estimate β_0, β_1, and γ_0?

4. Consider the model
$$y_t = \alpha x_t + u_t;\ t = 1, 2, \ldots, T$$
where $E(u_t) = 0$; $E(u_t^2) = \sigma^2 x_t$; $E(u_s u_t) = 0$ if $s \neq t$, for all s and t. The density function for u_t is
$$f(u_t) = (2\pi\sigma^2 x_t^2)^{-1/2} \exp\left[-\frac{1}{2\sigma^2}\left(\frac{u_t}{x_t}\right)^2\right].$$
Derive the maximum likelihood estimators of α and σ^2.

5. Consider the following data generating process
$$Y_t = Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3}$$
where ε_t is a white noise process. Assuming $-1 < \theta_2 < 0$ (and without taking a first difference), answer the following:
   (a) Find the h-step-ahead forecast for Y for $h = 1, 2, \ldots, k$
   (b) Find the h-step-ahead forecast error for Y for $h = 1, 2, \ldots, k$
   (c) Find the h-step-ahead forecast error variance for Y for $h = 1, 2, \ldots, k$

6. You have been given the task of estimating an earnings function using data from a data set on 2868 individuals. You have observations on:

EARN: log of earnings measured in dollars per hour
S: Years of schooling
ASVABC: the score on a test of cognitive ability
MALE: a dummy which = 1 if male, 0 if female
UNION: a dummy which = 1 if the individual belonged to a union in 2004, 0 otherwise.

You estimate four regressions using Ordinary Least Squares (OLS): regression (1) and (4) use the whole sample, regression (2) uses only those individuals who belonged to a union in 2004 and regression (3) uses only those individuals who did not belong to a union in 2004.

	Dependent	Variable: EARN		
	Whole sample (1)	Union Only (2)	Non-union only	Whole Sample (4)
S	0.066	0.028	0.070	0.066
	(0.004)	(0.012)	(0.005)	(0.004)
ASVABC	0.013	0.011	0.013	0.013
	(0.001)	(0.003)	(0.001)	(0.001)
MALE	0.214	0.286	0.199	0.209
	(0.017)	(0.049)	(0.018)	(0.017)
UNION	--	--	--	0.189
				(0.028)
Constant	0.819	1.545	0.750	0.803
	(0.055)	(0.164)	(0.058)	(0.055)
R^2	0.249	0.195	0.260	0.261
RSS	588.3	43.7	522.5	579.7
Sample size	2868	286	2582	2868

standard errors are in paranthesis.

(a) What, precisely, do the coefficients on MALE and UNION tell us?

(b) Using the above results analyse whether there is any difference between earnings for union and and non-union workers. Use whatever method you like but explain the method carefully, state all the assumptions necessary and discuss the limitations of your approach in answering the question of whether there is a difference..

(c) Another researcher points out that earnings are dependent on age and the length of time the individual has worked and suggests you include these two variables. You tell him you are concerned about multicollinearity if both these variables are included in this regression. What is the proble with multicollinearity and is it likely to occur in the revised equation? Explain.

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 (0.03)	−1.65 (0.26)	1.00
X_1	−0.13 (0.03)	−1.01 (0.28)	0.46
X_2	0.08 (0.03)	−0.74 (0.29)	0.66

(a) Consider an married man and a unmarried 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. 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, 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.