Econometrics Qualifier

You have 4 hours to complete this exam.

Please write legibly

Good Luck!

Part I

Qualified Exam Questions

(Total points: 50)

1. (10 points) Discuss in detail the White robust standard error in terms of:
   (1) The problem it solves.
   (2) The method it uses.

2. (10 points) Consider a linear regression model,

   $$y = x_1\beta_1 + x_1\beta_1 + u, \text{ where } E(X'u) \neq 0$$

   If $E(x_1) = E(x_2) = 0$, and $Cov(x_1, u) \neq 0$.

   Then: find out if $\hat{\beta}_{2OLS}$ is biased if:

   (1) $Cov(x_1, x_2) = 0$
   (2) $Cov(x_1, x_2) \neq 0$.

3. (15 points) For $T = 2$, consider a panel data model,

   $$y_{it} = x_{it}\beta + c_i + u_{it}, \quad t = 1, 2$$

   Let $\hat{\beta}_{FE}$ and $\hat{\beta}_{FD}$ denote the fixed effects and first difference estimators, respectively.

   (1) Show that $\hat{\beta}_{FE}$ and $\hat{\beta}_{FD}$ estimates are numerically identical.
   (2) Show that the error variance estimates from the FE and FD methods are numerically identical.
   (3) When $T > 2$, describe the difference between $\hat{\beta}_{FE}$ and $\hat{\beta}_{FD}$ if there is any.

4. (15 points) Consider a model,

   $$y = X\beta + u, \quad \text{where } cov(X, u) \neq 0.$$

   If we find a set of IVs such that $Cov(Z, u) = 0$, $Cov(X, Z) \neq 0$.

   (1) Show that the 2SLS is the best estimator among all linear IV estimators.
   (2) Is it necessary that the first stage estimator is consistent?
   (3) Discuss the advantages and/or disadvantages if we include Z in the main equation:

   $$y = X\beta + Z\gamma + u\ ?$$

Part II

Econometrics Qualifying Exam
Part II

[1] Consider a linear regression model $y_i = Z_i\beta + u_i$, $i=1,2,\cdots,n$, where Z_i is the $1\times K$ vector of strictly exogenous variables, and $u_i \sim i.i.d.N(0, \sigma^2)$. We wish to test a set of linear restrictions H_0: $R\beta - q = 0$ against H_1: $R\beta - q \neq 0$, where R is a $J\times K$ matrix of known constants and q is a $J\times 1$ vector of known constants. The restriction matrix R has a full row rank that is less than K, $rank(R) = J < K$.

(a) There are three classical methods of test: likelihood ratio (LR) test, Wald test and Lagrange multiplier (LM) test. We may interpret that each test compares the restricted and unrestricted estimators of a functionof parameters. Specify for each test what the function is and what its two estimators look like. For the LM test, you need to specify two alternative expressions. Note that this question is not asking you to write the details. Just brief description of the function and its estimators.

(b) The so called "Chow test" is typically applied to the test of stability of coefficients over different subsamples. It also compares two estimators. What are they? What is the test statistics and its null distribution for the hypothesis in this question?

(c) Hausman's test also compares two estimators. Give a general description of the two estimators that Hausman test compares. What are they for the hypothesis above in the current context of exogenous regressors Z_i?

(d) Now suppose that not all of the regressors in Z_i are exogenous, i.e., some of them are exogenous and some are correlated with the error term u_i. Let X_i be a $1\times L$ vector of variables that are exogenous with respect to u_i, where $L > K$. X_i may include the exogenous variables in Z_i. What is Hansen's J-statistic in this case? What hypothesis can we test by using it?

[2] Consider a latent variable regression equation $y_i^* = X_i\beta + u_i$, $i=1,2,\cdots,n$, where X_i is the $1\times K$ vector of strictly exogenous variables, and u_i is an i.i.d. normal random error term, $u_i \sim i.i.d.N(0, \sigma^2)$. We do not have the observations on all values of y_i^* : we have observations y_i

$$y_i = \begin{cases} y_i^* & if \;\; y_i^* > 0 \\ 0 & if \;\; y_i^* \leq 0 \end{cases}$$

(a) Derive $E(y_i|X_i)$ by using the hint below, and show why the OLS estimator $\hat{\beta}$ from a regression of y_i on X_i is biased. Which assumption of the Gauss-Markov theorem is violated?

(b) Specify the log likelihood function of y_i in terms of new parameters $\alpha \equiv \beta/\sigma$ and $\tau \equiv 1/\sigma$, and derive the first order conditions.

(c) Can you suggest an alternative estimation procedure other than the ML method? Briefly discribe. (Hint below may be useful)

(d) Present the decomposition of the marginal effect of a regressor X_{ik} on $E(y_i|X_i)$ and give a proper interpretation to each term.

(d) Suppose you run a regression by using a subsample of positive y_i^*. Show mathematically that the OLS estimator $\tilde{\beta}$ of such a regression is biased. Illustrate this bias by using a graph under the assumption that there is only one regressor (K=1). The graph should contains the mean line $E(y_i^* | X_i) = X_i\beta$ and show clearly which sample points will cause the regression line to be biased.

[Hint] Let $X \sim N(\mu, \sigma^2)$, $Z = (X-\mu)/\sigma$, $z=(x-\mu)/\sigma$ and $b=(c-\mu)/\sigma$. Let $\phi(z)$ and $\Phi(z)$ be the pdf and the cdf of Z. Then,

$$E(X|X>c) = \mu + \sigma\lambda^{+}(b), \quad \lambda^{+}(b) \equiv \frac{\phi(b)}{1-\Phi(b)}$$

$$E(X|X<c) = \mu + \sigma\lambda^{-}(b), \quad \lambda^{-} = -\frac{\phi(b)}{\Phi(b)}$$

[3] Consider a bivariate SVAR(1) and its reduced form VAR

(1) $\Gamma y_t = c_0 + C_1 y_{t-1} + u_t, \quad u_t \sim i.i.d.(0, \Sigma)$

(2) $y_t = a_0 + A_1 y_{t-1} + v_t, \quad v_t \sim i.i.d.(0, \Omega)$

where Γ, C_1 and A_1 are 2×2 matrices of coefficients, and Γ is a lower triangular matrix with principal diagonal elements being 1 ($\gamma_{ii}=1$), and Σ is a diagonal matrix.

(a) Show the restrictions on structural parameters such that the system is stable. You don't have to solve for the final set of conditions. Just derive the quadratic equation that includes the structural parameters and whose roots are the basis of finding the restrictions.

(b) Show that the model is just identified by showing that the only *admissible* transformation matrix is an identity matrix.

(c) Describe how you compute the structural parameters from the estimates of reduced form VAR.

(d) Suppose you estimate each structural equation by OLS. Are they consistent estimators? Are they the same as the estimators you find in (c)? Explain.