PhD/MA Econometrics Examination

January, 2016

Total Time: 8 hours

MA students are required to answer from A and B.

PhD students are required to answer from A, B, and C.

The answers should be presented in terms of equations, statistical details, and with necessary proofs and statistical deduction. Verbal and brief descriptive discussions will not be sufficient.

PART A

(Answer any TWO from Part A)

1. Fundamentals of OLS
   a. Write out the OLS equation in matrix form. Also, write out the matrices and state their dimensions.
   b. State the OLS assumptions in mathematical statements and in sentences (words).
   c. Show that the OLS estimator is BLUE and define BLUE. Show all parts: B, L, U, and E.
   d. What are the properties (hint: there are six) of the OLS estimator? State them in mathematics and words. Also, state any requirements which are necessary for these properties to hold.
   e. Given the properties in part d, what can you infer about the disturbances from the residuals?
   f. Write out a simple OLS model. Define your variables and describe how your model might meet or not meet all of the assumptions you stated above.

2. Variance of OLS
   a. What is the variance of the OLS estimator?
   b. Derive the estimator for the standard errors of the OLS estimator.
   c. Assuming σ^2 is unknown to the econometrician, what is the test statistic if the econometrician wants to test if a single element of the β vector is equal to zero or not?
   d. How is the test statistic in part c distributed and what are its degrees of freedom?
   e. Write out a simple OLS model. Define your variables and describe how you would use the test statistic in part d to interpret the output from your model.

3. Probability Theory and More

a. What distribution is below:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

b. Find the first and second moments of this distribution directly from the distribution itself.

c. Find the first and second moments using the moment generating function. Also, discuss if these moments are the same or different from the moments you found in **part d** and why or why not.

d. This distribution is one of two distributions with property called "memoryless." Prove that this distribution is memoryless.

e. Name a practical application of this distribution, i.e., where does it naturally occur?

f. Name the third and fourth central moments – *name them do not calculate them.*

g. The distribution above is a special case of another distribution. What distribution is it a special case of and exactly how is it a special case of that other distribution?

PART B

(Answer any TWO from Part B)

4. Maximum likelihood

   a. Let $X_1,\ldots,X_n$ be iid with pdf $f(x\,|\,\theta)=\frac{1}{\theta}e^{-x/\theta}$, $x\geq 0$, $\theta>0$

      i. What is the likelihood of observing your data (i.e., what is the likelihood function for your sample)?

      ii. Derive the log likelihood and score functions for estimating the parameter θ.

      iii. Derive the Maximum Likelihood Estimate for θ.

      iv. Derive the asymptotic variance for $\hat{\theta}_{MLE}$ using the information matrix method. *(Hint: Remember that $E[\theta]=\theta$ because it is a true parameter and thus just a number. And $E[x_i]=\bar{x}$. Use your answer in part iii to solve for $\bar{x}$ in terms of θ.*

5. Partitioned Regression

Consider the following model $Y = X_1\beta_1 + X_2\beta_2 + \varepsilon$, where X_1 is a matrix of k_1 variables and X_2 is a matrix of k_2 variables such that

$$X_1 = \begin{bmatrix} x_{11}^1 & x_{11}^2 & \cdots & x_{11}^{k_1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n}^1 & x_{1n}^2 & \cdots & x_{1n}^{k_1} \end{bmatrix}, \; X_2 = \begin{bmatrix} x_{21}^1 & x_{21}^2 & \cdots & x_{21}^{k_1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{2n}^1 & x_{2n}^2 & \cdots & x_{2n}^{k_1} \end{bmatrix}.$$

Denote b_1 and b_2 as the Ordinary Least Squares estimates for β_1 and β_2, respectively.

a. Derive the expression for the ordinary least squares estimator b_1 as a function of Y, X_1, X_2, and b_2 using the partitioned regression model.
    i. Suppose you only observe X_1 but not X_2. Thus you run the OLS model $Y = X_1\beta_1 + \varepsilon$.
    ii. Derive the expression for OLS estimate of b_1 that you would estimate under these conditions (i.e., what is the usual OLS estimator for b_1 when you regress Y on X_1).
    iii. If $Y = X_1\beta_1 + X_2\beta_2 + \varepsilon$ is the true model, give an expression for the amount b_1 (that you estimated in ii.1) is biased in this circumstance as a function of X_1, X_2, and b_2.
b. Now suppose you observe both X_1 and X_2. Derive the ordinary least squares estimator for and b_2 as a function of Y, X_1, X_2, and M_1 using the partitioned regression model. Where M_1 is the residual-maker matrix and $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.
c. Define the Frisch-Waugh Theorem and describe its intuition.
d. Under what conditions is the bias you solved for in part ii.2 equal to zero. What does this mean in the context of the Frisch-Waugh Theorem (i.e., what happens when you regress X_2 on X_1).

6. Suppose you want to estimate the effect of childbearing (motherhood status) on labor force earnings for a sample of women in the U.S. using the following model

(1) $$Y_i = \alpha + \beta D_i + \varepsilon_i ,$$

where Y_i is the labor market earnings of woman *i,* and D_i is a dummy variable equal to one if woman *i* has had at least one child. In this way you are hoping to estimate the average treatment effect of being a mother on female labor market earnings.

a. Which of the Ordinary Least Squares assumption is likely to fail when estimating this model? Explain why? What does the mean for your estimate of the average effect of motherhood on earnings?
b. What is the "Fundamental Problem of Causal Inference"?
c. Define Y_{1i} as the earnings of woman *i* if she is a mother and Y_{0i} as the earnings of that same woman *i* if she is not a mother. If motherhood status was randomly assigned across women in the population then the treatment effect of motherhood on earnings would be equal to the following

   (2) $$E_i[Y_{1i} - Y_{0i}] = E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0].$$

   However, motherhood status is not randomly assigned. Using the Potential Outcomes framework, decompose the expectation in (2) into the *"average treatment effect on the treated"* and *"selection bias".*
d. You decide to instrument for motherhood using the instrumental variable Z. Which two assumptions are necessary for this instrument to be valid?
e. Would the following variables be plausible instruments for motherhood. Explain why or why not.
   i. The quality of each woman's health insurance coverage
   ii. An indicator of whether or not the woman has experienced infertility
   iii. Regional differences in abortion laws (e.g. the oldest gestational age a woman can legally obtain an abortion in her region).
   iv. Availability of family planning in a local area.
   v. Number of siblings the woman has
   vi. The woman's marital status

PART C

(Answer any TWO from Part C)

7. A market survey is conducted to estimate consumers' preference for fresh food purchase behavior. They were presented with three options: A, B, and C, C being the status-quo. The data was collected on **income** and **age** of the consumers. The attributes considered were: 1) regular (Regular) versus organic (Organic) 2) the non-genetically altered (NGA) versus genetically altered (GA). In addition, the information about the **price ($3.50, $5.35,$ 7.10)** was also presented (/lb grocery bag on average) .

a. Using individual characters (Income and age) as the only determinants of the purchasing choices, set up a Random Utility Model (RUM) outlining the three indirect utility functions. Present the corresponding data table. You may assume that the income have different impact on the choice functions.
b. With the assumption of the extreme value distribution for the random error, present the three probability choices [P(A), P(B), and P(C)] and set up the log-likelihood function.
c. Now add the price and the other two attribute variables and present the three utility functions. Present the corresponding data table. You may assume that the income has the different impact on the utility functions, whereas the effect of price and age are the same.

8. Consider the following epidemiological model for the State of Texas, where the children's asthma rate (y – proportional to the children's population from 1-6 years of age) was expressed as a function of PM2.5 airborne pollution count (x). The data were collected for 254 counties for the year of 2010.

$$y = \frac{\emptyset x}{\delta + x} + \text{e}$$

which can be generically expressed as: $y = f(x, \emptyset, \delta) + e$

For simplicity, the subscript "t" is suppressed.

a. Using the generic expression 2, present the numerical estimation algorithm using the Newton Raphson optimization algorithm OR the maximum likelihood estimation method (e can be assumed to follow a normal distribution).
b. Discuss the method of deriving the variance-covariance matrix of the estimators.
c. The parameter $\emptyset$ is also known as the maximum y, thus it cannot be more than 1 (i.e., 100% asthma rate). With that in mind, can you conjecture some sort of graphical relation between y and x for the model given above?
d. Now, let's assume that you calculate the marginal impact of the pollution on the asthma rate. Explain the method of deriving the confidence level.

9. Consider the following VAR-model describing the relationship between the economic growth and the energy consumption, both expressed in terms of per capita for Canadian economy:

$$DInGPD(t) \qquad = C1 + \ \phi 11 * DInGDP(t-1) \quad + \phi 12 * DInEnCon(t-1) + u1(t)$$
$$DInEnCon(t) \ = C2 + \ \phi 21 * DInEnCons(t-1) \quad + \phi 22 * DInGDP(t-1) + u2(\text{t})$$

For simplicity, the model is expressed as a lag of order 1. The variables are logged, and are transformed as a first differencing (notice "D") to induce individual stationarity.

a. Derive the var-cov matrix of the error vector. Present in details the derivation, showing all the steps.
b. Why is it called a seemingly unrelated regression (SUR), and not a simultaneous model?
c. How would you rewrite the above model, if the two variables were cointegrated? What is the economic meaning of the cointegration between GDP and EnCon?
d. Present the GLS method and outline the estimation steps.

OR

Present the FIML likelihood procedure with the assumption of a multi-variate normal distribution.
e. Given the above VAR model, under what condition, would the GLS method be equal to the OLS? (Pick one of the two conditions and show the proof.)