Quantitative Methods Qualifying Exam

AAEC, UGA

May 17, 2018

Instructions:

- Answer all five questions.
- Begin each of the five questions with a new sheet of paper.
- Label your work with your ID number, NOT your name.
- To the extent possible, make your work easy to follow.

Good luck!!!

1. You have a simple, linear regression model, $y = X\beta + e$. You "know" that the error term is distributed according to the following process:

$$e_t = \gamma + \rho e_{t-1} + \lambda x_{3t} + u_t$$

where $u_t \sim N(0, \sigma^2)$, x_3 is the third variable in the X matrix, $t = 1, 2, ..., 50$ denotes time-ordered observations, and the regressor matrix X is full-column rank with 5 columns.

Show detailed steps that should be followed to arrive at consistent estimates of β.

2. Consider the following instrumental variable equation setup:

$$y_i = x_i\beta + \varepsilon_i$$

$$x_i = z_i\delta + \nu_i$$

where we have the following assumptions:

- The vector $(z_i, \nu_i, \varepsilon_i)$ is i.i.d.
- $E[z_i, \nu_i] = 0$
- $E[z_i^2] = \sigma_z^2 \in (0, \infty)$.

Define the IV estimator as:

$$b_{IV} = \left(\frac{1}{n}\sum_{i=1}^{n} z_i x_i\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} z_i y_i$$

(a) Is b_{IV} consistent? Prove your conclusion.

(b) Is b_{IV} precisely estimated? Prove your conclusion.

3. For every $i = 1, ..., n$, assume that:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Using $n = 100$ observations of $\{y_i, x_i\}_{i=1}^{n}$, a researcher (Hairy) estimates the following coefficients:

$$(b_0, b_1), (\hat{\sigma_{b_0}}, \hat{\sigma_{b_1}}), R^2, s^2$$

Suppose that another researcher (Dawg), is replicating Hairy's analysis. Dawg attempts to run the same regression that Hairy ran, but accidentally enters each observation twice. The new dataset has 200 observations, with the second $n = 100$ observations an exact repeat of the first 100 observations. Suppose that Dawg realizes his mistake, but no longer has access to a statistical package to run the regression again. How can Dawg compare his results to the results obtained by Hairy? Obtain closed form solutions for the following estimates that Dawg obtained, based on the esimates that Hairy obtained:

(a) (b_0, b_1) (i.e., the coefficient estimates),

(b) s^2 (i.e., the regression standard error),

(c) $(\hat{\sigma_{b_0}}, \hat{\sigma_{b_1}})$ (i.e., the standard errors of the coefficient estimates),

(d) R^2.

Note whether the estimates are the same or different. If they are different, state how they are different.

4. Consider the following regression model: $y_i = \beta x_{1i} + \gamma x_{2i} + \varepsilon_i$, where y, x_1, x_2, and ε are random variables (scalars) with the following properties:

- $E[x_{1i}] = 0$ and $E[x_{2i}] = 0$
- $0 < E[x_{1i}^2] < 1$ and $0 < E[x_{2i}^2] < 1$
- $E[x_{1i}x_{2i}] = 0$
- $E[\varepsilon_i | x_{1i}; x_{2i}] = 0$
- $E[\varepsilon_i^2 | x_{1i}; x_{2i}] = \sigma^2$

Suppose you have drawn an i.i.d. sample of size n from the model above.

(a) Consider $b_{OLS,full}$, the OLS estimator of β obtained from the regression of y_i on x_{1i} and x_{2i}. Prove that $b_{OLS,full}$ is a consistent estimator for β.

(b) Now consider $b_{OLS,small}$, the OLS estimator of β obtained from the regression of y_i on just x_{1i}. Prove that $b_{OLS,small}$ is a consistent estimator for β.

(c) Find the correct expressions for the variance of both $b_{OLS,full}$ and $b_{OLS,small}$.

(d) Show that $Var(b_{OLS,full}) < Var(b_{OLS,small})$.

(e) How should your result from part (d) be interpreted?

5. For decades, it had been assumed that addictions or habits are myopic in the sense that the consumer does not recognize the impact of his or her current decision on future health and preferences. The theory of rational addiction by Gary Becker and Kevin Murphy breaks from this tradition and hypothesizes that consumers can be forward-looking (i.e., rational) even when consuming addictive or habit-forming goods such as cigarettes and alcohol. This theory, if supported by data, has important implications for public policies such as tobacco control and war on drugs.

   Using cigarettes as an example, the model of rational addiction suggests the following demand relationship:

$$y_t = \alpha + \beta_1 y_{t+1} + \beta_2 p_t + \beta_3 y_{t-1} \tag{1}$$

   where y_t, y_{t+1}, and y_{t-1} are the number of cigarettes smoked at time t, $t+1$, and $t-1$, respectively; p_t is cigarette price at t. The parameters β_1 and β_3 measure the degree of forward- looking behavior and the degree of habits, respectively. There is no residual in the theoretical relationship shown in the model above because it assumes perfect foresight such that the level of y_{t+1} is known with certainty at time t. In the United States, state and federal excise taxes account for 43% of cigarette retail price. In addition, the level of state cigarette tax varies substantially across states and tax changes are known months before they are effective.

   (a) When applying equation 1 to longitudinal data on smoking, a residual term must be added, and y_{t+1} is no longer known with certainty at t. How would you estimate the parameters of equation 1?

   (b) Next, assume you estimate β_1 to be near-zero and statistically insignificant. Derive the short-run and long-run price elasticity of cigarette smoking.

   (c) How is the consistency of the OLS estimator dependent on the time-series property of the residual? Explain using math and intuition.