Econometrics Preliminary Exam
July 20, 2016

Read the instructions below before starting the exam:

- The preliminary exam consists of **three parts** (one for each quarter of Econ 480).
- You have 180 minutes (1 hour for each part) to solve this exam.
- This is a closed book and notes exam.
- You can use one page (both sides) of hand-written or typed notes.
- Remember to write your name on each examination booklet and on the space provided below.
- You may use a calculator if needed
- Please keep your written answers brief; be clear and to the point.
- We recommend that you circle or otherwise indicate your final answers.
- There is no choice, so remember to answer all questions.
- Good luck!

PLEASE WRITE EACH PART OF THE EXAM IN
A SEPARATE BLUE BOOK.

Name (First Last): ______________________________

2016 Econometrics Prelim, Part I (one hour, 60 points)

Let the grading of a pass-fail course be based on the student's total score on a set of examinations and problem sets. The range of feasible total scores is the real interval [0, 100].

Let y = a student's total score. Let w = 1 if the student passes the course and w = 0 if he fails. Let P(y, w) be the distribution of (y, w) in the population of students.

When answering all of the questions below, you know that grades are assigned as follows: w = 0 if $y \leq 50$ and w = 1 if y > 50. You also know that the fraction of the class who pass is P(w = 1) = 0.8.

1. (5 points) Given only the above information, what can you infer about E(y)?

2. (5 points) Given only the above information, what can you infer about M(y)?

3. (15 points) Suppose that P(y) is known to be the uniform distribution on the interval $[\alpha, \beta]$, for some values of α and β. What is the joint identification region for the pair of parameters α and β?

4. (20 points) Suppose that the student population is surveyed and asked to report their scores. Let z = 1 if a student reports his score on the survey and z = 0 otherwise. Assume that all reported scores are accurate. Suppose the survey reveals that P(z = 1) = 0.5 and that P(y|z = 1) is the uniform distribution on the interval [50, 90]. What can you infer about E(y)?

5. (15 points) Let x denote gender, with x = 1 if male and x = 0 if female. You are told that the gender composition of the class is as follows: P(x = 1) = 0.7 and P(x = 0) = 0.3. What can you infer about the fractions of males and females who pass the course; that is, about P(w = 1|x = 1) and P(w = 1|x = 0)]?

2016 Econometrics Prelim, Part II (1 hour, 60 points)

Do all problems, show all work, and define any symbols you use that are different from those in the problem statements.

Suppose the random variables Y, X, and ε are related by

$$Y = \beta X + \varepsilon,$$

where X is a scalar random variable, $E(X) = 0$, and X has as many higher order moments as you like. The unobserved random variable ε has mean 0, a finite but unknown variance, and as many higher order moments as you like. In addition $E(\varepsilon X) = \rho$ for some constant ρ that may not be zero. You have a random sample $\{Y_i, X_i : i = 1, ..., n\}$.

1. (6 points) Let $\hat{\beta}$ denote the ordinary least squares (OLS) estimate of β. What is the probability limit of $\hat{\beta}$ as $n \to \infty$?

2. (6 points) Under what conditions, if any, is $\hat{\beta}$ consistent for β? Prove that your answer is correct.

3. (6 points) If $\hat{\beta}$ is not consistent for β but you know ρ, is there a way to make a consistent estimator using only the data $\{Y_i, X_i\}$? If so, what is the estimator? Prove your answer is correct.

4. (9 points) If $\hat{\beta}$ is not consistent for β but you do not know ρ, is there a way to make a consistent estimator using only the data $\{Y_i, X_i\}$? If so, what is the estimator? Prove your answer is correct.

5. (12 points) The t-statistic for testing the hypothesis $H_0: \beta = \beta_0$ for some β_0 is

$$t = \frac{n^{1/2}(\hat{\beta} - \beta_0)}{s_n},$$

where s_n^2 is the variance of the asymptotic distribution of $n^{1/2}(\hat{\beta} - \beta)$. Find the asymptotic distribution of t if H_0 is true and $\rho = n^{-1/2}\rho_0$, where ρ_0 is a constant.

6. (12 points) Continue to assume that $\rho = n^{-1/2}\rho_0$. Suppose you reject $H_0: \beta = \beta_0$ at the nominal 0.05 level if $|t| > 1.96$. If $\rho_0 \neq 0$, is the asymptotic probability of rejecting a true H_0 larger or smaller than 0.05? Explain your answer.

7. (9 points) Now let H_0 be the hypothesis $H_0: \rho = 0$. Can you test this hypothesis using only the data $\{Y_i, X_i\}$? If yes, what is the test statistic? What is its asymptotic distribution under H_0?. If no, explain why.

Econometrics Preliminary Exam
Part III (60 points) - July, 2016

Question 1 30

Consider a Regression Discontinuity Design (RDD) like the one discussed in 480-3. There are potential outcomes $(Y(0), Y(1))$, a running variable Z, and a treatment assignment rule of the form

$$D = I\{Z \geq 0\} \ .$$

The singularity of the case here is that Z is a discrete random variable with five points of support, $z \in \mathcal{Z} \equiv \{-2, -1, 0, 1, 2\}$ such that $P\{Z = z\} > 0$ for all $z \in \mathcal{Z}$. Assume we observe an i.i.d. sample of n observations from the distribution of (Y, D, Z), where the observed outcome Y is given by

$$Y = DY(1) + (1 - D)Y(0) \ .$$

In addition, assume that $E[|Y(d)||Z = z] < \infty$ for $d \in \{0, 1\}$ and $z \in \mathcal{Z}$.

(a) (10 points) Show that the following average treatment effect is identified in this setting,

$$E[Y(1)|Z = 0] - E[Y(0)|Z = -1] \ .$$

(b) (10 points) The parameter of interest in RDD is usually the treatment effect "at" the cutoff, i.e.,

$$\theta = E[Y(1) - Y(0)|Z = 0] \ . \tag{1}$$

Is θ identified in this setting under the stated assumptions?

(c) (10 points) Suppose that $Y(0) = (1 - Z)^2 U$ where the unobserved random variable U satisfies $E[U|Z = z] = C$ for some (unknown) constant C and $z \in \{-1, 0\}$. Show that θ in (1) is identified.

Question 2 30

Let (Y, X, U) be a random vector where Y, X, and U take values in $\mathbf{R}$ and denote by P the distribution of (Y, X). Suppose that the causal model for Y states that, for $\beta = (\beta_0, \beta_1, \beta_2)' \in \mathbf{R}^3$,

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + U \ ,$$

where U and X are independent, $E[U] = 0$, $E[X] = 0$, $E[X^2] = 1$, $E[X^4] < \infty$, and $\beta_2 \neq 0$. The parameter of interest in this question is β_1. For the questions below, assume you have a random sample of n i.i.d. observations from P.

(a) (10 points) Consider the Least Squares (LS) projection of Y on X (i.e., a regression of Y on X and an intercept) and let $\hat{\beta}_{\text{ls}}$ denote the estimated projection slope. State the conditions (on P) under which $\hat{\beta}_{\text{ls}}$ is consistent for β_1. Show your results.

(b) Suppose that the conditions you derived in (a) do not hold, but that there is a binary instrument Z that is known to be independent of U. Consider the following estimator of β_1,

$$\hat{\beta}_{\text{new}} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

where

$$\bar{Y}_1 = \frac{1}{\sum_{i=1}^n Z_i} \sum_{i=1}^n Y_i Z_i \ , \quad \bar{Y}_0 = \frac{1}{\sum_{i=1}^n (1 - Z_i)} \sum_{i=1}^n Y_i (1 - Z_i) \ ,$$

and $\bar{X}_0$ and $\bar{X}_1$ are defined as $\bar{Y}_0$ and $\bar{Y}_1$ but with X_i replacing Y_i.

i. (10 points) Show that $\hat{\beta}_{\text{new}}$ is slope coefficient of the IV regression of Y on X and an intercept, using Z as an instrument.

ii. (10 points) Show that $\hat{\beta}_{\text{new}}$ is consistent for β_1 whenever

$$E[X|Z = 1] \neq 0 \text{ and } E[X^2|Z = 1] = 1 \ .$$

Note: depending on how do do the algebra you may alternatively arrive to

$$E[X|Z = 1] \neq E[X|Z = 0] \text{ and } E[X^2|Z = 1] = E[X^2|Z = 0] \ .$$