Problem Set 6
Econ 312, Spring 2019
James J. Heckman
Due May 16th, 2019

1. [5 pts] Consider a model

$$Y = X\beta + U \quad X \not\perp U$$

with instrument Z such

$$E(Z'X) \text{ has full rank}$$

$$E(Z'U) = 0.$$

Compare the consistency of 2 estimators.

   (a) IV (or 2 SLS) applied to sample of size N where $N \to \infty$

   (b) OLS for $N \to N^* \leq 10^{12}$

      TSLS for $N^* > 10^{12}$

2. [10 pts] Discuss the properties of the estimators obtained from the following specification-testing regression procedure for model

$$Y = X_1\beta_1 + X_2\beta_2 + U \qquad (*)$$

where (X_1, X_2) are scalars; $(X_1, X_2) \perp\!\!\!\perp U$

$|\text{Cov}(X_1, X_2)| > 0$; $(X_1, X_2, U) \sim N([\mu_1, \mu_2, 0]; \sum_{XU})$?

(a) Analyze the OLS estimators of β_1. From the full model $(*)$ $(X_1, X_2, U) \sim$ $N(\mu_1, \mu_2, 0; \sum_{XU})$.

(b) Derive OLS estimators of β_1 derived from the following procedure:

    i. Estimate (β_1, β_2) by OLS if $t_2 > 1.964$ (t_2 is t statistic for OLS estimate of β_2 when X_1 and X_2 are both entered).

    ii. Estimate β_1 as the value of π from the regression $Y = X_1\pi_1 + V$; $E(V) = 0$ if $t_2 < 1.964$.

This is called a *pretest estimator*. (Hint: See Bancroft, 1944.)

(c) What does this example say about the common practice of testing models using "specification tests"? How should we interpret the sampling distributions of estimates derived from pretest estimators?

3. [15 pts] Read Friedman's "Methodology of Positive Economics" (1953) and especially footnote 11.

(a) Define identification in general terms. Give an example.

(b) Compare the Cowles "structural approach" to the approach based on Friedman's methodology.

(c) If you use Friedman's approach, how do you learn from data? Compare it with the Cowles approach. Can you use the same data to build a model and test it? Can this be done rigorously? Does data reuse not risk pretest bias? Compare Bayesian, classical, and abductive approaches to learning from data. How does each approach deal with surprise? (Something unanticipated.) Distinguish be-

tween numerical surprises from model surprises, i.e., fundamentally new phenomena.

(d) Evaluate the following statement:

"Proper statistical inference requires that we specify hypotheses in advance of the data." No data snooping allowed.

4. [15 pts] Suppose you have a *finite* fixed size sample of a population of size N with some treated and some not. Treatment is randomly assigned. Each person has a (Y_0, Y_1) which is fixed. Only the sample of treatment assignments differ across individuals. They are determined by the toss of a coin. Let $X_i = 1$ is unit i assigned treatment. Unit i is assigned to control otherwise. You seek to estimate ATE by

$$\frac{1}{N} \sum_{i=1}^{N} (Y_{1i} - Y_{0i}).$$

(a) What is the standard error of the estimator

$$\frac{1}{N_1} \sum_{i=1}^{N} X_i Y_i - \frac{1}{N_0} \sum_{i=1}^{N} (1 - X_i) Y_i$$

where $N_1 = \sum_{i=1}^{N} X_0 N_0 = \sum_{i=1}^{N} (1 - X_i)$?

(b) Compare this sampling error with that from a model-based (super populative) approach. Compare the sources of variation in each approach.

5. [10 pts] Define the following concepts:

(a) Random sample.

(b) Truncated sample.

(c) Stratified random sample.

(d) Censored sample.

(e) Censored random variable.

(f) Choice based sample.

(g) Design-based inference vs model-based inference.

6. [15 pts] Define "causal parameter." Also define "structural parameter." Is a structural parameter a causal parameter? Is a causal parameter a structural parameter?

7. [20 pts] Discuss the credibility revolution in Econometrics (see Angrist and Pischke, 2010, *JEP*). What constitutes credibility? (See also companion paper by Keane.)

8. [10 pts] Suppose you have a set of 10 studies on the effect of repeal of drug laws on addiction.

(a) Define the properties of the vote counting method (i.e., what % of studies show a positive impact). What are its properties as the number of studies gets large?

(b) What would a meta-analyst report?

(c) Under what assumptions is the method of meta-analysis valid?

References

Angrist, J. D. and J.-S. Pischke (2010, Spring). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Bancroft, T. A. (1944, June). On biases in estimation due to the use of preliminary tests of significance. *The Annals of Mathematical Statistics 15*(2), 190–204.

Friedman, M. (1953). The methodology of positive economics. In M. Friedman (Ed.), *Essays in Positive Economics.* Chicago: University of Chicago Press.