

**WWS 508c, Spring 2008: Econometrics and Public Policy (Advanced)**

Monday/Wednesday, 9-10:30 AM, Bowl 002; precepts TBA.

Professor: Jesse Rothstein

jrothst@princeton.edu

Office hours: Tuesdays 2-4pm in 359 Wallace Hall, or by appointment

Preceptor: TBA

This course has three main goals. First, it will add to the toolkit that you began developing last semester, with several new techniques often used in empirical research and evaluation. Second, it will give you hands-on experience with real data and real analyses, aiming to add to your comfort level when it comes time for you to perform analyses of your own. Finally, and most importantly, it will try to give you some intuition about which tool to use in which situation, and attempt to persuade you to be skeptical of analyses that do not use appropriate tools for the topic at hand.

The course will be organized around problems that you might encounter in performing applied econometric analyses. These problems are not taken from any particular textbook. For reference, we will use the *Introductory Econometrics: A Modern Approach* by Jeffrey Wooldridge. In each unit, I have indicated relevant sections in this textbook. I have also indicated relevant sections in Stock and Watson's *Introduction to Econometrics*. These should be considered optional, unless otherwise noted, though if you find yourself having trouble with the math in Wooldridge, you may want to make a habit of reading the Stock and Watson selections for intuition before tackling the parallel treatment in Wooldridge. Finally, as we progress I will occasionally post articles from research journals on the course Blackboard site that may be of interest.

Computer exercises, both in class and as homework assignments, will be a key component of this course. You will need access to the STATA computer program, which is available on the PCs in the WWS computer clusters. I will make data sets available to you on the course web page.

If you are not yet familiar with Stata, there will be a somewhat steep learning curve. I recommend the book *Statistics with Stata*, by Lawrence C. Hamilton, as an introduction and tutorial reference. I have asked Labyrinth to order copies of this (and of Wooldridge), and I have put copies on reserve in Stokes Library. The Stata online help system can also be a useful reference, if you know what you are looking for. Early on in the course, you should familiarize yourself with the use, `insheet`, `infix`, and `save` commands; with commands for logging your work (`type help log`); with commands for creating new variables (`generate` and `egen`); and with commands to provide basic descriptions of your data (`describe`, `tabulate`, `list`, and `summarize`). You should do all of your Stata work—for the class, and in any later work—in “do” files, which make it much, much easier to keep track of what you’ve done and to reproduce your results. You will be required to submit your “do” and “log” files as part of your assignments.

### Who should take this class?

I believe that every MPP and MPA student should take one of the 508 courses—a familiarity with statistical analysis is invaluable in just about any policy context.

But the question of which flavor of 508 to take is a harder one. For the most part, the two flavors cover the same material, with different styles. 508c takes a more mathematical approach, while 508b is a bit lighter on the formal math. In my view, neither is “better” than the other; some people find math an extra hurdle, while others find it the easiest way to understand complex material. You should think about which type of person you are, and choose on this basis.

The mathematical level in this class will be higher relative to 508b than was 507c relative to 507b. That is, I expect that some students who took 507c last semester will want to take 508b this semester, but I expect very few who took 507b to enroll in 508c. You should be comfortable with the material in 507c—particularly the basic bivariate and multivariate regression models—and with differential calculus. If you aren’t, you will probably spend too much time trying to understand the mathematical details and will have trouble seeing the big picture. Since the big picture is the point, and the details are just a means to an end, this is a bad tradeoff.

I want to emphasize again that there is no value judgment in the choice of tracks. If you find math inherently confusing and non-intuitive, you will spend a great deal of time in 508c working on what seem like math problems, time that could be better spent developing the intuition in a better (for you) way in the B-track. You would almost certainly learn more in 508b. Every year, several students seem to make this mistake. If you think you might be one of them, please come talk to me early in the semester.

Please do not hesitate to ask me if you have any questions about your choice of courses.

### Course details

Listed below are twelve topics that I hope to cover during the term. I plan to spend about a week on each, moving a bit more quickly through the review units at the beginning to allow more time for the more difficult material at the end. We may not make it through the entire syllabus, however. At least part of our discussion of each topic will revolve around applications, either in class or on a problem set. Weekly precepts will provide reviews and extensions of the topics and techniques, and will also be a good place to get help with Stata.

Classroom participation is strongly encouraged, and will form 10% of your grade. There will also be homework assignments (approximately bi-weekly, totaling 35% of your grade), an in-class midterm (20%) and a take-home final exam (35%). I will post solution sets shortly after each homework assignment is due, so late assignments will not be accepted.

You are encouraged to work together on your homework, in groups of up to three. You may turn in a common copy of any Stata programs and log files, with all participants' names listed, but please write up any non-Stata answers separately. Whether you work in groups or not, you should make sure that you personally understand any work that you turn in—while it isn't expected that you did all of the programming on your own in the first place, by the time the assignment is handed in you should be able to do it without help.

## **Unit A: Review of the basic regression model**

### **I. Overview; Review of probability; Introduction to STATA**

- a. Several important rules about probability
- b. Introduction to STATA
  - Wooldridge, Sections 1.3, 1.4, and Appendices B and C
  - (Optional: Stock and Watson, Chapter 2, Section 15.2 and Appendix 16.1).

### **II. Univariate regression**

- a. Estimation of the mean as a regression problem
- b. Difference in means / t-test
  - W, Chapter 2, Sections 7.1-7.2
  - (SW, Chapter 4; Sections 3.4 and 15.4).
  - Krueger, Alan, "Experimental Estimates of Educational Production Functions." *Quarterly Journal of Economics* 114(2), May 1999, pp. 497-532.
  - ❖ Application: Class size effects in experiments and nonexperimental data

### **III. Multivariate regression**

- a. Short/long regression
  - W, Sections 3.0-3.2, 4.0-4.4 7.3
  - (SW, Chapter 5; Sections 16.1-16.5).
  - ❖ Application: Nonexperimental estimates of class size effects (cont.)

## **Unit B: What can go wrong with the basic regression model, what it can do to the estimates, and how one might deal with it**

- Recommended reading: SW, Chapter 7.

### **IV. Misspecification: $\text{corr}(X, u) \neq 0$**

- a. Omitted variables
- b. Errors in variables
- c. Endogeneity
  - W, Sections 3.3, 9.3, 9.4.
  - (SW, Sections 5.1, 5.11, 7.2)
  - Persico, Nicola, Andrew Postlewaite, and Dan Silverman, "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height."

*The Journal of Political Economy* 112(5), October 2004, pp. 1019-53.  
(Blackboard)

- Case, Anne, and Chris Paxson, “Stature and Status: Height, Ability, and Labor Market Outcomes.” Working paper, August 2006. (Blackboard)
- ❖ Application: Do tall people have higher IQs (part 3)?
- ❖ Application: Ability bias in estimating the returns to education

#### V. Things you can do when $\text{corr}(X, u) \neq 0$

- a. Instrumental variables
- b. Matching / propensity score
  - W, Sections 15.0-15.4.
  - (SW, Chapter 10.)
  - Angrist, Joshua D., and Alan B. Krueger, “Does Compulsory School Attendance Affect Schooling and Earnings?” *Quarterly Journal of Economics* 106(4), November 1991, pp. 979-1014. (Blackboard)
  - Angrist, Joshua D., and Victor Lavy. “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement.” *Quarterly Journal of Economics* 114(2), May 1999, pp. 533-575.
  - ❖ Application: Using geographic proximity to schools as an instrument for education

#### VI. Distribution of the residual: Independently and identically distributed?

- a. Efficiency of OLS with i.i.d. data
- b. Homoskedasticity vs. heteroskedasticity
- c. Weights
- d. Serial correlation in the basic regression model
  - W, Chapter 8
  - (SW, Sections 15.5 and 15.6.)
  - ❖ Application: The effect of student characteristics on school average outcomes.

### Unit C: Limited dependent variables

#### VII. Binary dependent variables

- a. Linear probability
- b. Probit/logit
  - Wooldridge, 17.0-17.1
  - (SW, Chapter 9)
  - ❖ Application: The effect of self employment on health insurance

#### VIII. Y is imperfectly observed.

- a. Censoring, truncation and sample selection.
- b. Tobit
- c. Heckman / control functions
  - Wooldridge, Sections 17.2, 17.4, 17.5
  - Clark, Melissa, Jesse Rothstein, and Diane Schanzenbach, “Selection Bias in College Admissions Test Scores.” Working paper. (Blackboard).
  - ❖ Application: Estimating school quality from average SAT scores

## **Unit D: Time series and panel data**

### **IX. Panel data techniques**

- a. Difference in differences
- b. Chi-square test
- c. Fixed effects & random effects
  - W, Chapters 13-14
  - Ashenfelter, Orley, and Michael Greenstone. “Using Mandated Speed Limits to Measure the Value of a Statistical Life.” *Journal of Political Economy*, February 2004. (Blackboard)
  - (SW, Chapter 8).
  - ❖ Application: Speed limits and traffic deaths

### **X. Time series**

- a. Autocorrelation and lagged dependent variables
  - W, Chapters 10-11
  - (SW, Chapters 12-13).
  - ❖ Application: Inflation, unemployment, and the Phillips curve

## **Unit E. Should we believe the results? How can we use them?**

### **XI. Research design / source of variation**

- a. Experiments
- b. Quasi-experiments
- c. Cross-sectional (and time series) inference
- d. Specification tests
  - Suggested reading: W, Section 7.6; SW, Sections 11.1-11.7—strongly recommended.
  - Goldin, Claudia, and Cecilia Rouse, “Orchestrating Impartiality: The Impact of Blind Auditions on Female Musicians.” *American Economic Review*, 90, no. 4 (September 2000): 715-741. (Blackboard)
  - Card, David, “The Impact of the Mariel Boatlift on the Miami Labor Market.” *Industrial and Labor Relations Review* 43 (2), 245-257, January 1990. (Blackboard).
  - Rothstein, Jesse, “Do Value-Added Models Add Value? Tracking, Fixed Effects, and Causal Inference.” Working paper, 2007. (Blackboard).