

WWS 508c Econometrics and Public Policy (Advanced)

Spring 2005

Monday/Wednesday, 9-10:30 AM, Robertson 005; precepts TBA.

Professor: Jesse Rothstein (jrothst@princeton.edu; OH Monday, 1-3, or by appt.)

Preceptor: Piotr Elias

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 teach you when each tool is called for, 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. There are a few topics that Wooldridge does not cover to my taste. Also, some of you may be more comfortable with a somewhat less mathematical presentation. I have also indicated relevant sections in Stock and Watson, which you used last semester. These should be considered optional, unless otherwise noted. 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 7*, by Lawrence C. Hamilton, as an introduction and tutorial reference. I have asked the U-Store to order copies, and I have put a copy on reserve in the WWS 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* 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.

This class will be more difficult 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. Please do not hesitate to ask me if you have any questions about your choice of courses.

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 weekly or bi-weekly homework assignments (40%), an in-class midterm (20%) and a take-home final exam (30%). 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. I will post solution sets shortly after each problem set is due, so late assignments will not be accepted.

Unit A: Basic regression model

I. Overview; Review of probability; Introduction to STATA

- a. Several important rules about probability
- b. Introduction to STATA
 - Suggested readings: W, Sections 1.3, 1.4, and Appendices B and C (SW, 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
 - Suggested readings: W, Chapter 2, Sections 7.1-7.2 (SW, Chapter 4; Sections 3.4 and 15.4).
 - ❖ Application: Election returns and Palm Beach County

III. Multivariate regression

- a. Short/long regression
 - Suggested readings: W, Sections 3.0-3.2, 4.0-4.4 7.3 (SW, Chapter 5; Sections 16.1-16.5).
 - ❖ Application: Do workers who use computers have higher wages?

Unit B: Improved inference when the classical assumptions aren't satisfied

IV. 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
 - Suggested readings: W, Chapter 8 (SW, Sections 15.5 and 15.6.)
 - ❖ Application: The effect of student characteristics on school average outcomes.
 - ❖ Application: The effect of layoff announcements on stock prices.
 - ❖ Application: Estimating the NAIRU (non-accelerating inflation rate of unemployment).

V. Binary dependent variables

- a. Linear probability
- b. Maximum likelihood
- c. Probit/logit
 - Suggested readings: Wooldridge, 17.0-17.1 (SW, Chapter 9).
 - ❖ Application: The effect of self employment on health insurance
 - ❖ Application: Women's labor force participation

Unit C: What can go wrong, what it can do to the estimates, and how one might deal with it

- Suggested reading: SW, Chapter 7.

VI. Misspecification: $\text{corr}(\mathbf{X}, \mathbf{u}) \neq 0$

- a. Omitted variables
- b. Errors in variables
- c. Endogeneity
 - Suggested readings: W, Sections 3.3, 9.3, 9.4. (SW, Sections 5.1, 5.11, 7.2)
 - ❖ Application: Ability bias in estimating the returns to education
 - ❖ Application: Income as an LHS variable, annual vs. permanent
 - ❖ Application: Income as an RHS variable, annual vs. permanent

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

- a. Instrumental variables
- b. Matching / propensity score
 - Suggested readings: W, Sections 15.0-15.4. (SW, Chapter 10.)
 - ❖ Application: Using geographic proximity to schools as an instrument for education
 - ❖ Application: Using education as an instrument for mismeasured income
 - ❖ Application: Evaluation of job training programs
 - ❖ Application: Cigarette smoking and infant birth weight

VIII. Misspecification: \mathbf{Y} is imperfectly observed.

- a. Censoring, truncation and sample selection.
- b. Tobit
- c. Heckman / control functions
 - Suggested readings: Wooldridge, Sections 17.2, 17.4, 17.5
 - ❖ Application: How long do transplanted kidneys last?
 - ❖ Application: Married women's labor supply (hours of work)

- ❖ Application: The effect of education on married women's earnings

IX. Panel data techniques

- a. Difference in differences
- b. Chi-square test
- c. Fixed effects & random effects
 - Suggested readings: W, Chapters 13-14 (SW, Chapter 8).
 - ❖ Application: Minimum wages and employment in NJ and PA
 - ❖ Application: Workers compensation laws and duration of benefit receipt
 - ❖ Application: Drunk driving laws and traffic deaths

X. Nonlinearity

- a. Polynomials
- b. Nonparametrics
- c. Regression discontinuity
 - Suggested readings: W, Section 6.2. (SW, Chapter 6).
 - ❖ Application: Effect of education on earnings, with controls for ability
 - ❖ Application: Effect of incumbency on election probabilities

Unit D. 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
 - Suggested reading: W, Section 7.6; SW, Sections 11.1-11.7—strongly recommended.
 - ❖ Application: Effects of vouchers on student test scores
 - ❖ Application: Using the Mariel boatlift to estimate the effect of immigration on the local labor market

XII. External validity

- a. Out of sample prediction
- b. Heterogeneous treatment effects
 - ❖ Application: Using quarter of birth as an instrument for education