

Political Science 509: The Linear Model

Emory University, Spring 2006

Tarbutton Hall 120A

TuTh 1:00p-2:15p

January 19, 2006

Professor: Eric Reinhardt
Office: 329 [Tarbutton Hall](#)
Office hours: Th 9:30a-11:30a, & by appt.
Phone: 404-727-4977
Email: erein@emory.edu
My home page: <http://userwww.service.emory.edu/~erein/>
Course home page: <http://userwww.service.emory.edu/~erein/courses/pols509/>
Course directory: [\\soc-sci.ss.emory.edu\polisci\polsclass-lab\pols_509\](http://soc-sci.ss.emory.edu/polisci/polsclass-lab/pols_509/)

Course Description & Objectives

This course covers basic techniques in quantitative political analysis. It introduces students to widely-used procedures for regression analysis, and provides intuitive, applied, and formal foundations for regression and more advanced methods treated later in the course sequence. Unlike POLS 508, this course will use rudimentary calculus and matrix algebra rather intensively, but prior knowledge of such methods (beyond basic high school algebra) is not obligatory. Rather, the course includes brief primers on those mathematical techniques. In terms of foundations, the course also covers some probability theory and the properties of estimators. As its core, the course uses these foundations to examine multivariate ordinary least-squares regression. The course covers model assumptions and techniques for addressing violations of those assumptions (e.g., heteroskedasticity, autocorrelation, multicollinearity), as well as issues of model specification, functional forms, measurement error, and endogeneity.

Requirements

Grades in the course will be based on the following items:

- **50%** — Homework. Eleven equally-weighted assignments, typically with a week from distribution to due date. Turn these in by the start of class. Unlike the 508 assignments, these will be graded; they may not be pass/fail. You may *talk* about the problems with your fellow students, but *do not copy* another person's homework; your submission must contain your own work. Please hand in homeworks in hard copy, rather than emailing them, unless otherwise approved.
- **50%** — Methods paper, 15-20 pages. This paper will demonstrate your technical mastery of the practical aspects of OLS regression in the context of a specific research problem of your own formulation. Choose a topic, develop a hypothesis, and test it quantitatively. The only twist here is that your dependent variable should be continuous rather than discrete, so that it is suitable for the kinds of techniques we will be learning in class. The format should be similar to a "research note" in *APSR* or *JOP*, but with greater emphasis on the technical details. A one-page synopsis describing the hypothesis to be tested and the dataset to be used is due on Th, Feb 16, at the start of class. **Final version due Wed, May 2.** Paper [guidelines](#) may be found on the course web page.

Course Policies

Late assignments will be penalized. Each day the assignment is late will result in a drop of a letter grade, e.g., A to B, etc.

Reading Materials

Statistics texts are expensive, but they are the kind of book that you'll find yourself referring to frequently throughout your graduate years and beyond. We will use a variety of complementary texts, listed below. You may come to prefer one over the other two. Greene is hardest but most thorough and detailed; Gujarati is relatively broad but shallow in some places; Wooldridge is strong on intuition and explanations but not as formal as the other two. Feel free to buy and use whichever of these three you like; they can be found at the Druid Hills bookstore in Emory Village. Wherever they occur on the required reading list, they are substitutes. But reading more than one can help, often a great deal.

- Early on in the course we will use the monograph, Timothy M. Hagle, *Basic Math for Social Scientists: Concepts* (Sage, 1995). This is required wherever it is listed.
- Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, 2nd ed. (Thomson South-Western, 2003). 3rd edition may also be available; page numbers for both editions are listed below.
- William Greene, *Econometric Analysis*, 5th ed. (Prentice Hall, 2003). A favorite among political scientists; has many practical applications and variations of the main models and tests (with somewhat more emphasis on cross-sectional sorts of problems), but uses matrix algebra as the primary exposition device, and does not spell out every single step in some cases. A difficult first-read.
- Damodar N. Gujarati, *Basic Econometrics*, 4th ed. (New York: McGraw-Hill, 2003).

The syllabus also occasionally lists optional reading selections from Peter Kennedy, *A Guide to Econometrics*, 4th ed. (MIT, 1998); John Fox, *Applied Regression Analysis, Linear Models, and Related Methods* (Sage, 1997); or other sources. Check [NetLibrary](#) for free online availability for some of these. (You can also navigate to this by searching for them through EUCLID.) Any other reading on the syllabus can be found in the course's Soc-Sci folder.

Like POLS 508, this course will rely on *Stata* as our chief statistical software. Just as before, you can buy your own personal *Stata* license if you wish. Call the *Stata* Corporation at 800-782-8272, saying that you are part of the "GradPlan III" for Emory University, if you want to place an order; or go to <http://www.stata.com/info/order/new/edu/gradplans/gp3-order.html>. *Stata* 8.0 is installed on Tarbutton's local network server, Soc-Sci, at <\\soc-sci.ss.emory.edu/polisci/apps/stata>.

Course Outline

Jan 19 (Th) &

- Jan 24 (Tu): (1) Introduction. Course administration. Math primer I: notation, sets, functions, limits, differential calculus.
- Hagle, 1-54.
 - Wooldridge, 675-695 [707-727]¹.
 - Optional: Fox, 521-523.

¹ The main page numbers listed for Wooldridge are for his 2nd edition; the pages in [] brackets are for the 3rd edition.

Jan 26 (Th) &

Feb 2 (Th): (2) Math primer II: more differential calculus, optimization, integral calculus.

- Hagle, 58-71.
- Optional: Fox, 535-539.
- ⇒ Homework # 1 distributed.

Jan 31 (Tu): **No class.** (Reinhardt presenting research at World Bank and American Enterprise Institute.)

Feb 7 (Tu) &

Feb 9 (Th): (3) Math primer III: probability theory and distributions, estimators and their properties.

- Wooldridge, 696-775 [728-777].
- Greene, 845-890, esp. 885-890.
- Gujarati, 870-905.
- Optional: [Kennedy](#), 1-29.
- Optional: Fox, 540-566.
- ⇒ Homework # 2 distributed.

Feb 14 (Tu) &

Feb 16 (Th): (4) The regression model. Estimation and inference.

- Wooldridge, 21-89, 95-187, 197-198 [23-94, 99-197].
- Hagle, 54-56.
- Gujarati, 58-81, 100-114, 119-147, 202-215, 229-232, 294-296 (plus 114-118 optional).
- Optional: Greene, 7-71.
- Optional: Fox, 85-94, 112-125.
- ⇒ Homework # 3 distributed.
- ⇒ Paper **proposal due** Th Feb 16.

Feb 21 (Tu): (5) Model fit and outliers. Dummy variables, interactions. Simulating substantive effects and other estimated quantities.

- Wooldridge, 182-256 [192-261].
- Gujarati, 81-87, 217-223, 248-273, 297-333.
- Optional: Gary King, Michael Tomz, and Jason Wittenberg, "Making the Most of Statistical Analyses: Improving Interpretation and Presentation," *American Journal of Political Science* 44:2 (April 2000), 341-355.
- Optional: Fox, 267-287, 135-154.
- Optional: [Kennedy](#), 221-232.
- Optional: Michael S. Lewis-Beck and Andrew Skalaban, "When to Use R-Squared," *The Political Methodologist* 3:2 (1990), 9-11.
- Optional: Gary King, "When *Not* to Use R-Squared," *The Political Methodologist* 3:2 (1990), 11-12.
- Optional: Robert C. Luskin, "R-Squared Encore," *The Political Methodologist* 4:1 (1991), 21-23.
- Optional: Robert J. Friedrich, "[In Defense of Multiplicative Terms in Multiple Regression Equations](#)," *American Journal of Political Science* 26 (November 1982), 797-833.
- ⇒ Homework # 4 distributed.

Feb 23 (Th) &

Feb 28 (Tu): (6) Math primer IV: matrix algebra.

- Hagle, 71-95.
 - Wooldridge, 776-786 [808-818].
 - Gujarati, 913-925.
 - Greene, 803-845.
 - Optional: Fox, 524-534.
- ⇒ Homework # 5 distributed.

Mar 2 (Th) &

Mar 7 (Tu): (7) The regression model in matrix form. Estimation and inference.

- Wooldridge, 787-801 [819-833].
 - Gujarati, 926-949.
 - Greene, 7-71.
 - Optional: Fox, 204-218, 221-235.
- ⇒ Homework # 6 distributed.

Mar 9 (Th): (8) Multicollinearity.

- Wooldridge, 96-100 [101-105].
 - Gujarati, 335-375.
 - Greene, 56-59.
 - Optional: Fox, 337-366.
 - Optional: [Kennedy](#), 183-193.
- ⇒ Homework # 7 distributed.

Mar 14 (Tu) &

Mar 16 (Th): **No class.** Spring Break.

Mar 21 (Tu) &

Mar 23 (Th): (9) Heteroskedasticity. The problem & diagnosis. GLS and robust SEs.

- Wooldridge, 257-288 [271-298].
 - Gujarati, 387-440 (856-862 optional).
 - Greene, 215-249, 198-211.
 - Optional: Fox, 301-309, 320-321, 326-328.
 - Optional: George W. Downs and David M. Roche, "[Interpreting Heteroscedasticity](#)," *American Journal of Political Science* 23:4 (November 1979), 816-828.
 - Optional: Jeffrey B. Lewis and Drew A. Linzer, "Estimating Regression Models in Which the Dependent Variable is Based on Estimates," *Political Analysis* 13 (2005), 345-364.
- ⇒ Homework # 8 distributed.

Mar 28 (Tu) &

Mar 30 (Th): (10) Autocorrelation. The problem, diagnosis, techniques, GLS and robust SEs.

- Wooldridge, 391-424 [412-442, optional 363-373].
 - Gujarati, 441-505.
 - Greene, 250-282.
 - Optional: Fox, 369-385.
 - Optional: [Kennedy](#), 121-126.
 - Optional: Frank R. Baumgartner, Bryan D. Jones, and Michael C. MacLeod, "[The Evolution of Legislative Jurisdictions](#)," *Journal of Politics* 62:2 (May 2000), 321-349.
- ⇒ Homework # 9 distributed.

- Apr 4 (Tu): (11) Model specification.
- Gujarati, 506-523, 529-548, 556-560.
 - Greene, 116-161.
 - Wooldridge, 89-95, 100-101, 142-148, 198-200, 289-322 [94-99, 192-205, 304-334].
 - Optional: Fox, 235-239, 59-74, 309-317.
 - Optional: [Kennedy](#), 88-99.
- Apr 6 (Th) &
Apr 11 (Tu) &
Apr 13 (Th): (12) Techniques for panel data.
- Wooldridge, 426-483 [448-476, 485-501].
 - Gujarati, 636-652, 273-279.
 - Greene, 283-338.
- ⇒ Homework # 10 distributed.
- Apr 18 (Tu): (13) Selection bias and the Heckman selection model.
- Wooldridge, 585-592 [616-623].
 - Greene, 780-789.
 - Kevin B. Grier, Michael C. Munger, and Brian E. Roberts, “The Determinants of Industry Political Activity, 1978-1986,” *American Political Science Review* 88:4 (December 1994), 911-926.
 - Optional: Anne E. Sartori, “An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions,” *Political Analysis* 11 (2003), 111-138.
- ⇒ Homework # 11 distributed.
- Apr 20 (Th): **No class.** (Reinhardt at Midwest conference.)
- Apr 25 (Tu) &
Apr 27 (Th): (14) Measurement error & endogeneity. Instrumental variables, concepts, adequacy, testing. Two-stage least squares.
- Wooldridge, 295-309, 484-524, optional 525-552 [510-574].
 - Greene, 83-90, 378-400.
 - Gujarati, 524-529, 717-778.
 - Optional: Fox, 130-133.
 - Optional: Steven D. Levitt, “[Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime](#),” *American Economic Review* 87:3 (June 1997), 270-290.
- May 3 (W): **Paper due** by 1:00p.