

Political Science 509: The Linear Model

Emory University, Spring 2007

Tarbutton Hall 120A

TuTh 10:00p-11:15p

22 January 2007

Professor: Kyle Beardsley
Office: 317 Tarbutton Hall
Office hours: Th 1:00-3:00, & by appt.
Phone: 404-727-7506
Email: kyle.beardsley@emory.edu
My home page: <http://userwww.service.emory.edu/~kbeards/>
Course directory: \\as.eu.emory.edu\POLShare\POLClass-Lab\POL-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. This course builds on the prior efforts of Chris Zorn (a former faculty member now at South Carolina) and Eric Reinhardt.

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 15, at the start of class. **Final version due Thursday, May 3.** Paper guidelines can be found on the course directory.

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; Greene and Wooldridge should be at the campus bookstore. Wherever they occur on the required reading list, they are substitutes. But reading more than one can help, often a great deal. If you choose to go only with Greene, you will likely find yourself in over your head. I recommend the combination of Greene and Wooldridge.

- 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*, 3rd ed. (Thomson South-Western, 2003). This text offers clear step-by-step descriptions and instructions. If you find yourself completely lost on a given topic, Wooldridge is often the best place to look for answers.
- 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>.

Course Outline

Jan 18 (Th) &

- Jan 23 (Tu): (1) Introduction. Course administration. Math primer I: notation, sets, functions, limits, differential calculus.
- Hagle, 1-54.
 - Wooldridge, 707-727.
 - Optional: Fox, 521-523.

Jan 25 (Th) &

- Jan 30 (Tu): (2) Math primer II: More differential calculus, optimization, integral calculus.
- Hagle, 58-71.
 - Optional: Fox, 535-539.
- ⇒ Homework # 1 distributed.
- Feb 1 (Th) &
Feb 6 (Tu): (3) Math primer III: Probability theory and distributions, estimators and their properties.
- Wooldridge, 728-777.
 - Greene, 845-890, esp. 885-890.
 - Gujarati, 870-905.
 - Optional: [Kennedy](#), 1-29.
 - Optional: Fox, 540-566.
- ⇒ Homework # 2 distributed.
- Feb 8 (Th) &
Feb 13 (Tu): (4) The regression model: estimation and inference.
- Wooldridge, 23-94, 99-197, optional 276-7.
 - Hagle, 54-56.
 - Gujarati, 58-81, 100-114, 119-147, 175-178, 202-215, 229-232, 294-296 (plus 114-118 optional).
 - Optional: Greene, 7-71, optional 940.
 - Optional: Fox, 85-94, 112-125.
- ⇒ Homework # 3 distributed.
- ⇒ Paper **proposal due** Th Feb 15.
- Feb 15 (Th): (5) Model fit and outliers. Dummy variables, interactions. Simulating substantive effects and other estimated quantities.
- Wooldridge, 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 20 (Tu) &
Feb 22 (Th): (6) Math primer IV: Matrix algebra.
- Hagle, 71-95.
 - Wooldridge, 808-818.
 - Gujarati, 913-925.

- Greene, 803-845.
 - Optional: Fox, 524-534.
- ⇒ Homework # 5 distributed.

Feb 27 (Tu) &

Mar 6 (Tu): (7) The regression model in matrix form. Estimation and inference.

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

Mar 1 (Th): No Class; I will be presenting at the ISA meeting in Chicago

Mar 8 (Th): (8) Multicollinearity.

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

Mar 13 (Tu) &

Mar 15 (Th): **No class.** Spring Break.

Mar 20 (Tu) &

Mar 22 (Th): (9) Heteroskedasticity. The problem & diagnosis. GLS and robust SEs.

- Wooldridge, 271-298.
 - Gujarati, 387-440, optional 856-862.
 - 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 27 (Tu) &

Mar 29 (Th) &

Apr 3 (Tu): (10) Autocorrelation and Intro to Time Series. The problem, diagnosis, techniques, GLS and robust SEs. Non-stationarity and unit roots. Spatial autocorrelation.

- Wooldridge, 412-442, 363-373, 380-405, 639-645
- Gujarati, 441-505.
- Greene, 250-282, 631-649.
- Nathaniel Beck, Kristian Skredre Gleditsch and Kyle Beardsley, "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy," *International Studies Quarterly* 50:1 (2006), 27-44.
- Optional: Jeffrey M. Wooldridge, "Cluster-sample Methods in Applied Econometrics: An Extended Analysis," Working Paper (June 2006).

- 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 5 (Th): (11) Model specification.
- Gujarati, 506-523, 529-548, 556-560.
 - Greene, 116-161.
 - Wooldridge, 94-99, 192-205, 304-334.
 - Optional: Fox, 235-239, 59-74, 309-317.
 - Optional: [Kennedy](#), 88-99.

- Apr 10 (Tu) &
Apr 12 (Th) &
Apr 17 (Tu): (12) Techniques for panel data.
- Wooldridge, 448-476, 485-501.
 - Gujarati, 636-652, 273-279.
 - Greene, 283-338.
 - Optional: Jeffrey M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*; Cambridge, MA: MIT Press (2002), 282-283.
- ⇒ Homework # 10 distributed.

- Apr 19 (Th): (13) Selection bias and the Heckman selection model.
- Wooldridge, 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 24 (Tu) &
Apr 26 (Th): (14) Measurement error & endogeneity. Instrumental variables, concepts, adequacy, testing. Two-stage least squares.
- Wooldridge, 310-325, 510-551, optional 552-581.
 - 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 (Th): **Paper due by 5:00p.**