

Political Science 503
Quantitative Methods
Spring 2013
Mondays and Tuesdays, 9:00-10:15 AM
Rosenkranz 06

Professor Thad Dunning
420 Rosenkranz
thad.dunning@yale.edu
Office Hours: Tuesdays 2:00-3:00 PM or by appt.
Teaching Fellow: Dawn Teele (Office hours TBA)

This course is the second semester in the graduate quantitative methods sequence in the political science department. We begin with the potential outcomes model, considering several topics including causal and statistical inference; sampling distributions and standard errors; classical hypothesis testing, as well as permutation tests; and the estimation of causal effects in the presence of non-compliance with treatment assignment.

Next, we turn to linear regression and several extensions. We begin with the mechanics of regression in both bivariate and multivariate forms. We then discuss the linear regression model, where observed data are viewed as realizations of random variables. Finally, we consider causal inference under this model. Model specification is a central area of concern. We will discuss several techniques for addressing limited breakdowns in modeling assumptions, such as instrumental-variables least-squares regression. It is important to master technique, and this course will help you do so; it is even more important to understand the core assumptions.

After the spring break, we turn to special topics in the design and analysis of experimental and observational studies, such as difference-in-differences designs; matching; and natural experiments, including regression-discontinuity designs. We also cover computer simulations, including the bootstrap. The role of strong research design in generating valid causal inferences is given special emphasis throughout the course.

Each week's reading typically includes methods readings as well as one or more applied article relevant to the week's focus. Weekly sections will focus especially on building computing skills.

The course assumes that students have command of the material covered in PLSC 500, including basic knowledge of probability and linear regression. Matrix algebra and calculus are helpful; however, we will be covering those topics as we go.

The URL for the course website is <http://www.thaddunning.com/teaching/plsc-503>. I will try to post lecture slides for the week to the course website in advance of each week's classes. Viewing or printing the slides in advance may help you focus on the concepts rather than on taking extensive notes.

Problem sets: Working exercises is crucial for consolidating understanding. Problem sets will be handed out most Tuesdays, and each problem set will be due the Monday after it is handed out. The problem sets will combine theoretical or conceptual problems with computer exercises.

I will ask you to form small groups to work together on the problem sets. However, you should always try to work out all the answers yourself before meeting with your group. No late problem sets will be accepted. The teaching fellow Dawn Teele and I will be available to help inside and outside of class.

Exams: There will also be a take-home midterm and a take-home final exam, to be taken individually. Grades will be based on exams (40% for the midterm, 50% for the final), and participation in lectures and sections and completion of problem sets (10%).

Section: There will be a two-hour discussion section led by Dawn Teele. This is a chance for you to reinforce your understanding of the lecture material and, especially, to learn to use statistical packages. Replication of published research results using publically available data sets and computer simulation routines such as the bootstrap will be covered in lectures, sections, and problem sets.

Statistical Packages: You can use a statistical package of your choosing to work the computer exercises on the problem sets. Stata is still probably the most widely used package in political science, and this is mostly what Dawn will be teaching. However, the open-source package R is increasingly used, and learning R is a good idea too. Matlab is also nice option (though few political scientists seem to use it); the lab exercises in Freedman (2009) are written using Matlab. Unfortunately, different social scientists work (and post their replication data and code) using different software packages. So for awhile, you may be taking Spanish and French at the same time... Ultimately, you'll probably pick one package and use it for most of your work, but it is good to have familiarity with a few of the most widely used ones.

Readings (Required; listed roughly in order of frequency of use):

Freedman, David A. 2009. *Statistical Models: Theory and Practice*. New York: Cambridge University Press, 2nd edition.

Gerber, Alan, and Donald Green. 2012. *Field Experiments: Design, Analysis, Interpretation*.

Freedman, David, Robert Pisani, and Roger Purves. 2007. *Statistics*. New York: W.W. Norton & Company, Fourth Edition.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge University Press.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*.

Freedman, David and David Lane. 1981. *Mathematical Methods in Statistics: A Workbook*. New York: W.W. Norton & Company.

Recommended:

Freedman, David A. (David Collier, Jasjeet S. Sekhon, and Philip B. Stark, eds.) 2009. *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*. New York: Cambridge University Press.

Greene, William H. 2011. *Econometric Analysis*. Prentice Hall, 7th edition (earlier editions are cheaper and are fine).

Simon, Carl P. and Lawrence Blume. 1994. *Mathematics for Economists*. W.W. Norton & Company.

Wooldridge, Jeffrey M. 2009. *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning, Fourth edition.

Some notes on the readings: You should buy all the books if you can; their value to you over the years will more than compensate for the hefty initial investment. Be sure to buy the second (2009) edition of the Freedman's *Statistical Models*. Freedman and Lane (1981) is now out of print, but when last I checked, many cheap, used copies were available through Amazon; the book might seem a little quaint, but it is quite helpful. It may be substantially cheaper to buy an earlier edition of the Wooldridge book, and that should be fine. Greene is another standard and oft-cited econometric reference. Simon & Blume (1994) is on the list because it provides an introduction to matrix algebra (e.g., Chapters 6-11); it may also be useful to have on your shelf for other purposes. If you prefer to have another matrix algebra book, see Freedman (2009) for references.

Course schedule and readings

N.B. I have tentatively organized the material as noted below, with one slide presentation for each week, but we may need to adjust the timing as we go.

It is extremely important to keep up with the reading and to get on top of technical material as soon as it is introduced; the course will prove difficult if you do not do this.

Jan. 14-15: The potential outcomes model and causal inference

- Gerber and Green 2012, Chapter 2
- Dunning 2012, Chapters 1 and 5 (p. 105-21 only)
- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102 (1): 33-48.
- Recommended: Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945-960.

Jan. 21-22: Sampling distributions, statistical inference, and hypothesis testing

- Gerber and Green 2012, Chapter 3
- Dunning 2012, Chapter 6, pp. 165-73, 186-95.
- Freedman, Pisani and Purves 2007, Chapters 5, 13-14, 16-18, 20-21, 26-27
- Dunning, Thad, and Lauren Harrison. 2010. "Cross-Cutting Cleavages and Ethnic Voting: An Experimental Study of Cousinage in Mali." *American Political Science Review* 104 (1): 21-39.

[NOTE: We will need to meet on Jan. 21 (MLK Jr. day), because Thad is away on Friday, Jan. 18 when Monday classes meet]

Jan. 28-29: Non-compliance; instrumental variables

- Dunning, 2012, Chapter 5 (pp. 135-57)
- Gerber and Green, 2012, Chapters 5-6
- Angrist, Joshua D. 1990. "Lifetime Earnings and the Vietnam-Era Draft Lottery: Evidence from Social Security Administration Records." *American Economic Review* 80 (3): 1284-86.
- Recommended: Robert S. Erikson and Laura Stoker. 2011. "Caught in the Draft: The Effects of Vietnam Draft Lottery Status on Political Attitudes." *American Political Science Review* 105 (2): 221-37.
- Recommended: David Clingingsmith, Asim Ijaz Khwaja, and Michael Kremer. 2009. "Estimating the Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering." *Quarterly Journal of Economics* 124 (3): 1133-1170.

Feb. 4-5: The mechanics of bivariate regression

- Freedman, Pisani, and Purves 2007, Chapters 8-12.
- Freedman and Lane 1981, all, but especially pp. 48-86.
- Freedman 2009, Chapters 1-2.
- Douglas A. Hibbs. 1978. "On the Political Economy of Long-Run Trends in Strike Activity." *British Journal of Political Science* 8: 153-75.

Feb. 11-12: The mechanics of multiple regression

- Freedman (2009), Chapter 3-4
- Angrist and Pischke, Chapters 2-3.
- Recommended: Simon and Blume 1994, Chapters 6-11.
- Yule, G. Udny. 1899. "An Investigation Into the Causes of Changes in Pauperism in England, Chiefly During the Last Two Intercensal Decades." *Journal of the Royal Statistical Society* 62: 249-95.

Feb. 18-19: The linear regression model: causal/statistical inference; hypothesis testing

- Freedman 2009, Chapters 5 and 6.
- Gerber and Green 2012, Chapter 10

- Gibson, James L. 1988. "Political Intolerance and Political Repression During the McCarthy Red Scare." *American Political Science Review* 82 (2): 511-29 (included in Freedman 2009, pp. 315-342).

Feb. 25-26: Instrumental variables: the IVLS regression model

- Freedman (2009), Chapter 9.
- Dunning 2012, Chapter 9, pp. 269-278, 285-6.
- Angrist and Pischke (2009), Chapter 4.
- Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 122: 725-53.
- Recommended: Sovey, Allison J. and Donald P. Green. 2009. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55: 188-200.

March 4-5: Bootstrapping a regression model; clustering

- Freedman (2009), Chapter 8
- Dunning 2012, Chapter 6 (pp. 175-86, 195-201)
- Beck, Neal and Jonathan Katz. 1995. "What to Do (and Not To Do) With Time-Series Cross-Section Data." *American Political Science Review* 89 (3): 634-647.

Take-home Midterm Exam, March 7-8.

March 8-24– Spring Break

March 25-26: Difference-in-Differences; Times-Series Cross-Sectional Models

- Angrist and Pischke 2009, Chapter 5
- Wooldridge 2009, Chapters 10-12, 13-14.
- Green and Gerber, 2012, Chapter 4.
- Donald T. Campbell and H. Laurence Ross. 1968. "The Connecticut Crackdown on Speeding: Time-Series Data in Quasi-Experimental Analysis." *Law & Society Review* 3(1): 33-54.

April 1-2: Matching; missing data and experimental attrition

- Angrist and Pischke (2009), Section 3.3
- Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2006. "Comparing Experimental and Matching Methods using a Large-Scale Voter Mobilization Experiment." *Political Analysis* 14: 1-36.
- Sekhon, Jasjeet S. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." *Annual Review of Political Science* 12: 487-508.
- Gilligan, Michael J. and Ernest J. Sergenti. 2008. "Do UN Interventions Cause Peace? Using Matching to Improve Causal Inference." *Quarterly Journal of Political Science* 3 (2): 89-122.
- Green and Gerber, Chapter 7.

April 8-9: Natural experiments, RDD, and Tests of Design; Interference between units

- Dunning, 2012, Chapters 5 (pp. 121-136, 158-160), 7, and 8. Read Chapters 2-4 for background.
- Angrist and Pischke 2009, Chapter 6.
- F. Daniel Hidalgo, 2012. "Digital Democratization? The Political Consequences of Electronic Voting in Brazil and India." Manuscript, Department of Political Science, MIT.
- Recommended: McCrary, Justin. 2007. "Testing for Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 615-635.

April 15-16: More Tests of Design; Interference between Units

- Gerber and Green, 2012, Chapter 8.
- Nickerson, David. 2008. "Is Voting Contagious? Evidence from Two Field Experiments." *American Political Science Review* 102: 49-57.
- Caughey, Devin M. and Jasjeet S. Sekhon. 2011. "Elections and the Regression-Discontinuity Design: Lessons from Close U.S. House Races, 1942–2008." *Political Analysis* 19 (4): 385-408.

April 22-23: Wrap up: model- and. design-based inference; multi-method research

- Freedman, David. 2009. Chapters 1-3, 20 in Freedman, David A. (David Collier, Jasjeet S. Sekhon, and Philip B. Stark, eds.), *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*. New York: Cambridge University Press.
- Dunning, 2012, Chapters 11-12
- Freedman 2009, Chapter 10
- Gerber and Green 2012, Chapters 11-12.

Prepare for Final Exam

Take-home Final Exam, April 29-30