Syllabus
GOV 2002 Topics in Political Methodology

Professors: Matthew Blackwell and Arthur Spirling
Fall Semester 2014

Class Room
K354 CGIS
10:00-12:00
Thursday

Office
Matthew Blackwell
Room K-305, CGIS
Phone: (617) 496-2426
Email:mblackwell@gov.harvard.edu

Arthur Spirling
Room K-436, CGIS
Phone: (617) 495-3953
Email: aspirling@gov.harvard.edu

Preliminaries

Overview and Class Goals

This course is designed to introduce students to several broad areas of modern statistics and to illustrate how ideas and methods from each of these areas can be fruitfully applied to political science. The course is not designed to provide a comprehensive treatment of any of these topics. Instead, it is designed to introduce students to key ideas and to provide enough background knowledge so that students can pursue a more detailed study of some topics on their own.

Prerequisites

Students should have taken GOV 2000 and GOV 2001 or the equivalent courses. Some students without this background may still be able to do very well in this course. Students who have not taken courses similar to GOV 2000 and GOV 2001 should talk to the instructors to get a sense of how prepared they are for the course.

Class Requirements

Students are expected to attend all meetings. The course is formally structured as a lecture, followed by structured discussion and presentation. We (the instructors) would prefer to have the course operate as an ongoing conversation among class participants. Many of the topics are either open areas of research and/or somewhat controversial and thus are well-suited to detailed discussion. We encourage students to ask questions and spark discussion. We are certainly open to students suggesting directions for the class, and suggesting readings for the various weeks.

Assessment will be based on presentation and participation (20% of final grade), homework(s) (20% of final grade) and a final project (60% of final grade). The final project has a 'prospectus' component (4 pages) which are to be handed in to the instructors the class before the Thanksgiving holiday period begins (November 22). The prospectus should lay out the paper: the idea (perhaps the substantive question of interest), the techniques to be used, where the data will be found, the results expected. The final projects are designed to allow students to develop a deeper understanding of one or more of the topics introduced in the course. Possible final projects include (but are not limited to):

- Writing a paper that applies some of the methods discussed in class to a substantive question with real data.
- Building an original dataset from textual data and conducting some preliminary analysis.
- Developing a new model or procedure.
- Proving a conjecture put forth by the instructors.

Final projects need to be approved by both instructors by December 13.

We will not give incompletes in this course.

Each week, two students will present readings from the syllabus in an APSA-style presentation: 12–15 minutes, with slides. Students should be sure to critique the works, thinking about its strengths, weaknesses, and how it could/should be applied in political science. The class meets on Thursdays: students should email the instructors by Sunday should they face particular problems understanding/interpreting the papers under study.

Computation

We will use the R computing language in this course. The R language is completely open-source, so you can download it for free from http://www.r-project.org/. R will also be installed at HMDC and on the FAS computing cluster.

We hope to have classes on `Python` also, which will be introduced in our weeks on text analysis. *Prior knowledge of Python is not necessary for this course.*

Course Website

The course website is located at the following URL: http://isites.harvard.edu/k98475. This site will provide homework assignments, datasets, and supplementary materials.

Office Hours and Availability

Professors Blackwell and Spirling will have office hours by appointment this semester. If you have questions about the course material, computational issues, or other course-related issues please do not hesitate to set up an appointment with the Professors.

Required Books

Angrist, Joshua and Pischke, Jörn-Steffen. 2009. *Mostly Harmless Econometrics.*

Optional Books

Bird, Steven; Ewan Klein; and Edward Loper. 2007. *Natural Language Processing in Python.* http://nltk.sourceforge.net/index.php/Book

Bishop, Christopher M. 2006. *Pattern Recognition and Machine Learning.* New York: Springer.

Gelman, Andrew et al. 2004. *Bayesian Data Analysis.* New York: Chapman & Hall.

Gill, Jeff. 2002, 2007. *Bayesian Methods: A Social and Behavioral Sciences Approach.* (1st or 2nd Ed). New York: Chapman & Hall

Hastie, Trevor; Robert Tibshirani; and Jerome Friedman. 2001. *The Elements of Statistical Learning*. New York: Springer.

Hernán, Miguel and Robins, James. 2011. *Causal Inference (Part I)*. http://www.hsph.harvard.edu/faculty/miguel-hernan/causal-inference-book/

Manning, Christopher D. and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: The MIT Press.

Morgan, Stephen L. and Winship, Christopher. 2007. *Counterfactuals and Causal Inference*. New York: Cambridge University Press.

Pearl, Judea 2000. *Causality*. Cambridge: Cambridge University Press.

Rosenbaum, Paul. 1995. *Observational Studies*. New York: Springer.

Simonoff, Jeffrey S. 1996. *Smoothing Methods in Statistics*. New York: Springer.

Spirtes, Peter; Glymour, Clark; and Scheines, Richard. 1993. *Causation, Prediction, and Search*. New York: Springer Verlag.

Venables, W.N. and B.D. Ripley. 2002. *Modern Applied Statistics with S-PLUS*. New York: Springer.

Wasserman, Larry. 2006. *All of Nonparametric Statistics*. New York: Springer.

Preliminary Schedule

The course outline below is a ***preliminary*** schedule. We will adjust the amount of time spent on each topic to meet student interest and may well cut some of the topics below to make room for additional material. Students should let the instructors know what they are most interested in.

The "Required Reading" should be completed prior to lecture in a given week. "Optional Reading" is directly related to the lecture for the week in question but is not required. The "Optional Reading" typically provides more detail on some of the themes topics covered in lecture. The "Background Reading" provides supplementary material that may be helpful but is not necessarily directly related to what is covered in lecture.

1 Sept 11. Bayesian Analysis I: Philosophy and Approach

Topics covered

- nature of probability and uncertainty
- Bayesian setup: priors and posterior inference
- Hypothesis testing in Bayesian framework

Required Reading

Presentation Brad Efron "Why Isn't Everyone a Bayesian?" The American Statistician, Vol. 40, No. 1 (Feb., 1986) [include following discussion of Efron's article]

Optional Reading

Gill, Ch 1–7

2 Sept 18. Bayesian Analysis II: Estimation via MCMC

Topics covered

- Monte Carlo methods/integration
- sampling approaches
- Markov Chain Monte Carlo
- convergence and practical concerns

Required Reading

Gill Ch 8, 9

Presentation Bruce Western and Simon Jackman. 1994. "Bayesian Inference for Comparative Research" The American Political Science Review , Vol. 88, No. 2 , pp. 412-423

Presentation Simon Jackman. 2000. "Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo" American Journal of Political Science, Vol. 44, No. 2., pp. 375-404.

Optional Reading

Robert, Christian P and George Casella. 2004. "Monte Carlo Statistical Methods", 2nd Ed. Springer. [esp ch 6 and 7]

Gilks, W, S. Richardson and David Spiegelhalter. 1995."Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics", Chapman & Hall.

3 Sept 25. Causal Inference I: Causal Identification and Estimation under Conditional Independence

Topics covered

- Potential outcomes framework
- Selection on observables and the back-door criterion
- Regression, matching, weighting, and doubly robust estimators

Required Reading

Hernán & Robins, Ch. 1, 7.

Imbens & Rubin, Ch. 4 (skim), 6-7.

Presentation Angrist & Pischke, Ch 1-3.3.

Presentation Imbens, G. 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review". Review of Economics and Statistics 86 (1): 4–29.

Optional Reading

Angrist & Pischke, Ch 3.3-3.4.

4 Oct 2. Causal Inference II: Instrumental Variables

Topics covered

- Instrumental variables
- Local average treatment effects and compliance

Required Reading

Angrist and Pischke: Chapter 4

Presentation Angrist, Joshua D., Imbens, Guido W., & Rubin, Donald B. (1996). "Identification of Causal Effects Using Instrumental Variables (with discussion)." *Journal of the American Statistical Association*, 91, 444455.

Presentation Guido Imbens. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of Economic Literature*. 48(2): 399-423 & Deaton, Angus. 2010. "Instruments, Randomization, and Learning About Development." *Journal of Economic Literature*. 48(2): 42455.

Optional Reading

Abadie, Alberto 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics*. 113 (2003) 231-263.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak." *Journal of the American Statistical Association*. 90(430), 443.

Kern and Hainmueller, 2009. "Opium for the Masses: How Foreign Free Media Can Stabilize Authoritarian Regimes." *Political Analysis*.

5 Oct 9. Roll Call Analysis

Topics Covered

- summarizing data: clustering, PCA
- item response theory: Rasch model and others
- spatial voting
- likelihood-based approaches: NOMINATE
- Bayesian approaches, identification
- practical problems in roll call analysis: party discipline

Required Reading

Presentation Clinton, Joshua D., Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." American Political Science Review 98:355–370

Presentation Poole, Keith and Howard Rosenthal. 1997 "Congress: A Political-Economic History of Roll Call Voting" New York: Oxford University Press. Ch 2,3.

6 Oct 16. Causal Inference III: Regression Discontinuity

Topics covered

- Sharp and fuzzy regression discontinuity designs
- Sorting at the discontinuity

Required Reading

Angrist and Pischke: Chapter 6

Presentation Lee, David S. 2008. "Randomized Experiments from Non-random Selection in U.S. House Elections". *Journal of Econometrics*. Volume 142, Issue 2, Pages 675-697.

Presentation Eggers, Andrew C., Olle Folke, Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder, Jr. "On The Validity Of The Regression Discontinuity Design For Estimating Electoral Effects: New Evidence From Over 40,000 Close Races." Forthcoming, *American Journal of Political Science*.

Optional Reading

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142: 615-35. (Part of special issue on RDD, all of which is of interest.)

Hahn, J., P. Todd and W. van der Klaauw (2001), "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design," *Econometrica*, vol. 69: 201-209.

7 Oct 23. Causal Inference IV: Repeated Measurements over Time

Topics covered

- Random and fixed effects
- Differences-in-differences estimators

Required Reading

Angrist & Pischke: Chapter 5

Presentation Abadie, A., Diamond, A., & Hainmueller, J. (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of Californias Tobacco Control Program." *Journal of the American Statistical Association*, 105(490), 493505.

Presentation Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119(1): 249-75.

Optional Reading

Sobel, Michael E. (2012). "Does Marriage Boost Mens Wages?: Identification of Treatment Effects in Fixed Effects Regression Models for Panel Data." *Journal of the American Statistical Association*, 107(498), 521529.

Imai, Kosuke, and In Song Kim. 2012. "On the Use of Linear Fixed Effects Regression Models for Causal Inference." Working paper. http://imai.princeton.edu/research/files/FEmatch.pdf

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies* 72(1): 119.

Card, D. and A. B. Krueger (1994), "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, vol. 84, 772-793.

Blackwell, Matthew (2013). "A Framework for Dynamic Causal Inference in Political Science." *American Journal of Political Science*, 57(2), 504520.

8 Oct 30. Strategic Interaction in Political Science

Topics covered

- strategic models
- comparative statics analysis
- strategic selection bias
- belief updating

Required Reading

Presentation Signorino, Curtis. 1999. Strategic Interaction and the Statistical Analysis of International Conflict. American Political Science Review. 93(2):279-98.

Presentation Bas Muhammet, Curtis Signorino, and Robert Walker. 2008. Statistical Backwards Induction: A Simple Method for Estimating Recursive Strategic Models. Political Analysis 16 (1): 21-40.

Optional Reading

Signorino, Curtis, 2003. Structure and Uncertainty in Discrete Choice Models. Political Analysis . 11(4)

Carrubba, C. J., A. Yuen, and C. Zorn. 2007. In defense of comparative statics: Specifying empirical tests of models of strategic interaction. Political Analysis 15:46582

Signorino, Curtis 2007. On Formal Theory and Statistical Methods: A Response to Carrubba, Yuen, and Zorn. Political Analysis 15:483-501

Leemann, Lucas. 2014. Strategy and Sample Selection: A Strategic Selection Estimator. Political Analysis (2014) pp. 124

9 Nov 6. Texts as Data

Topics covered

- n-gram Representations
- The "Bag of Words" Assumption
- Stemming
- The Vector Space Model and Term Weighting
- Unsupervised, Supervised, and Semi-Supervised Learning

Required Reading

Presentation Quinn, Kevin M.; Burt L. Monroe; Michael Colaresi; Michael H. Crespin; and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." American Journal of Political Science, 54, 209–228.

Presentation Hopkins, Daniel, and Gary King. "A Method of Automated Nonparametric Content Analysis for Social Science." American Journal of Political Science 54, no. 1 (2010): 229-247.

Optional Reading

Manning and Schütze. 1999. *Foundations of Statistical Natural Language Processing*, Chapters 4, 5, 6, 14, 15, and 16.

Background Reading

Blei, David M. and John D. Lafferty. 2007. "A Correlated Topic Model of *Science*." forthcoming in *Annals of Applied Statistics*.
http://www.cs.princeton.edu/∼blei/papers/BleiLafferty2007.pdf

Thomas, Matt; Bo Pang; and Lillian Lee. 2006. "Get Out the Vote: Determining Support or Opposition from Congressional Floor-Debate Transcripts".
http://www.cs.cornell.edu/home/llee/papers/tpl-convote.dec06.pdf

10 Nov 13. Text II: Regular Expressions and Parsing

Guest lecture/hands on session.

11 Nov 20. Text III: Using Python for Web Scraping

Guest lecture/hands on session

Required Reading

Bird, Steven; Ewan Klein; and Edward Loper. 2007. *Natural Language Processing in Python*.
http://nltk.sourceforge.net/index.php/Book
Chapters 2 and 3.

12 Dec 4. Experiments (Enos)

Topics covered

- Randomization Inference
- Experimental Design
- Experimental Artifacts

Required Reading

Gerber, Alan, and Donald P. Green. (2012). Field Experiments. Norton: New York. Chapters 2 and 3.

Presentation Imai, Kosuke, Gary King, and Clayton Nall. (2009). The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation" (with discussions and rejoinder), Statistical Science, Vol. 24, No. 1 (February), pp. 29-53.

Presentation Rosenthal, Robert, and Donald B. Rubin. "Interpersonal expectancy effects: The first 345 studies." Behavioral and Brain Sciences 1.03 (1978): 377-386.

Optional Reading

McDermott, Rose. (2011) "Internal and External Validity" in Cambridge Handbook of Experimental Political Science, Eds. Druckman et al.

13 Dec 11. Student Presentations of Paper Topics: Group Critique