

Syllabus

GOV 1000/2000/2000e Quantitative Methods for Political Science I

Professor: Adam Glynn

TFs: Richard Nielsen and Brandon Stewart
Fall Semester 2010

Class Room

Room N-354, CGIS
2-4 PM
Tuesdays

Office

Adam Glynn
Room N-305, CGIS
Phone: (617) 496-2426
Email: aglynn@fas.harvard.edu

Richard Nielsen
Email: rnielsen@fas.harvard.edu

Brandon Stewart
Email: bstewart@fas.harvard.edu

Preliminaries

Overview and Class Goals

How can we detect voting irregularities? Is Islam incompatible with women's rights or open government? What causes individuals to vote? In what sense (if any) does democracy (or trade) reduce the probability of war? Quantitative political scientists address these questions and many others by using and developing statistical methods that are informed by theories in political science. In this course, we provide an introduction to the tools used in basic quantitative political methodology. The first four weeks of the course cover introductory univariate statistics, while the remainder of the course focuses on linear regression models. Furthermore, the principles learned in this course provide a foundation for the future study of more advanced topics in quantitative political methodology.

While the tools of statistical inference are worth studying in their own right, the primary goal of this course is to provide graduate students (and some undergraduates) with the necessary skills to critically read, interpret, and replicate the quantitative content of many political science articles. As such, the statistical methods covered in this course will be presented within the context of a number of articles. Throughout the term, we will reanalyze the data and revisit the conclusions from

- "Islam and Authoritarianism," by M. Steven Fish, *World Politics*, 2002.
- "Racial Prejudice and Attitudes Toward Affirmative Action," by James H. Kuklinski, Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers, *The American Journal of Political Science*, 1997.
- "Measuring Political Preferences," by Lee Epstein and Carol Mershon, *American Journal of Political Science*, 1996.
- "Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues," by Ebonya L. Washington, *The American Economic Review*, 2008.
- "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment," by Alan S. Gerber, Donald P. Green, and Christopher W. Larimer, *The American Political Science Review*, 2008.

- “Law and Data: The Butterfly Ballot Episode,” by Henry E. Brady, Michael C. Herron, Walter R. Mebane Jr., Jasjeet Sekhon, Kenneth W. Shotts, Jonathan Wand, *PS: Political Science and Politics*, 2001.
- “The Classical Liberals Were Right: Democracy, Interdependence, and Conflict 1950-1985,” by John R. Oneal and Bruce M. Russett, *International Studies Quarterly*, 1997.
- “MPs for Sale? Returns to Office in Postwar British Politics,” by Andrew Eggers and Jens Hainmueller, *The American Political Science Review*, 2009.

Class Requirements

Grades will be based on

- weekly homework assignments (50 % of final grade)
- a midterm exam (10 % of final grade)
- a cumulative take-home final exam (30 % of final grade)
- participation and presentation (10 % of final grade).

The weekly homework assignments will consist of analytical problems, computer work, and data analysis. The 1000, 2000, and 2000e sections of the course will have different homework assignments, but there will be some overlap. For all sections, the homework will be assigned on Tuesday night and due the following Monday at 5:00pm. The homework write-up must be word processed (MS Word is fine), with tables and figures incorporated in the document. No late homework will be accepted except in the case of a documented emergency. All sufficiently attempted homework¹ will be graded on a (+,√,-) scale, and may be re-written and re-graded once. The re-writes will be due on Monday at 5:00pm, one week after the assignment is returned. I encourage students to work together on the assignments, but you should write your own solutions (this includes computer work), and I strongly suggest that you make a solo effort at all the problems before consulting others. The take-home final exam will be handed out one week before the last Friday of reading period. Ten percent of the grade will be awarded for class participation and for the quality of presentation on the homework assignments and the exams.

You should bring pencil and paper to class because I will give some short in-class assignments.

I will not give incompletes in this course.

The 1000, 2000, and 2000e Course Numbers

Gov 2000 is designed for students who already have some background in statistics/mathematics/computing, or for beginners who are looking for a challenge. Students taking this section of the course will learn to be flexible data analysts, capable of tailoring standard methods to the unique specifications of each task. As such, these students will be asked in the problem sets to write/adjust the code necessary to replicate and critique results from the literature. This section of the course will be taught within the R statistical computing environment.

Gov 2000e is designed for students with a limited background in statistics/mathematics/computing. Students in this section of the course will focus on the analysis and critique portions of the assignments. This section of the course will be taught with the Stata statistical software, and the students will be provided with the additional code necessary to replicate and critique results from the literature.

Gov 1000 is intended for undergraduate students.

¹The instructor will determine sufficiency in borderline cases. All sufficiently attempted homework will be typed and well organized with all problems attempted.

Discussion Sections

There will be two discussion sections for this course. The 2000 section will discuss the concepts and R code needed to complete the Gov 2000 homework. The 2000e/1000 section will discuss the concepts and Stata procedures needed to complete the Gov 2000e and 1000 homework.

Course Website

The course website is located at the following URL:

<http://www.courses.fas.harvard.edu/2281>. This site will provide homework assignments, datasets, and supplementary materials.

Course Mailing List

Information about the course mailing list is available at the course web-site under the “WWW Links” section. All students should subscribe to the list. This an ideal forum for posting questions regarding the course material and/or computing. I encourage students to reply to each other’s questions, and a student’s respectful and constructive participation on the mailing list will count toward his/her class participation grade.

Office Hours and Availability

My office hours will be Monday from 5:00pm - 7:00pm this semester in room N-401 of CGIS. At these office hours, I will answer questions about submitted homework problems as well as other questions regarding the course. I am also available by appointment.

If you have a general question, you can also send it to the course mailing list. This is almost always the fastest way to get an answer. However, you can also email me directly at aglynn@fas.harvard.edu. If the question is of general interest, I will forward the question and my answer to the list. Make sure to tell me explicitly in your email if you would like to stay anonymous.

Required Books

(ALZ) Ashenfelter, Orley, Levine, Philip, and Zimmerman, David. 2003. *Statistics and Econometrics: Methods and Applications*. John Wiley & Sons.

(GH) Gelman, Andrew and Hill, Jennifer. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Note: Most of the required material from ALZ can also be found in the Wooldridge text listed first in the optional books section.

Optional Books

The following books are optional but may prove useful to students looking for additional coverage of some of the course topics.

Wooldridge, Jeffrey. 2000. *Introductory Econometrics*. New York: South-Western.

Fox, John. 1997. *Applied Regression Analysis, Linear Models, and Related Methods*. Thousand Oaks, CA: Sage.

Fox, John. 2002. *An R and S-PLUS Companion to Applied Regression*. Thousand Oaks, CA: Sage.

Gill, Jeff. 2006. *Essential Mathematics for Political and Social Research*. 1st Edition. 2nd printing. New York: Cambridge University Press.

Weisberg, Sanford. 2005. *Applied Linear Regression*. 3rd Edition. Hoboken, NJ: John Wiley.

- Freedman, David; Robert Pisani; and Roger Purves. 1998. *Statistics*. 3rd Edition. New York: Norton.
- Agresti, Alan and Finlay, Barbara. 1997. *Statistical Methods for the Social Sciences* Upper Saddle River, NJ: Prentice Hall.
- Cleveland, William S. 1993. *Visualizing Data*. Summit, NJ: Hobart Press.
- Simon, Carl and Blume, Lawrence. 1994. *Mathematics for Economists*. New York: Norton.
- Kennedy, Peter. 2003. *A Guide to Econometrics*. 5th Edition. Malden. Blackwell.
- Wonnacott, Thomas H. and Ronald J. Wonnacott. 1990. *Introductory Statistics*. 5th Edition. New York: Wiley.
- Venables, W.N. and B.D. Ripley. 2002. *Modern Applied Statistics with S-PLUS*. New York: Springer
- Gonick, Larry and Smith, Woollcott. 1993. *The Cartoon Guide to Statistics* New York: Harper.
- Tufte, Edward. 2001. *The Visual Display of Quantitative Information, 2nd Edition*. Cheshire, CN: Graphics Press.

Preliminary Schedule

The “Required Reading” should be completed prior to the lecture for which it is assigned. The “Suggested Reading” presents alternative presentations of the required material. The “Optional Reading” provides supplementary material that may be helpful or interesting. *This schedule is subject to change.*

1 Introduction - Sept. 7

Topics Covered

- Overview and Course Requirements
- Examples of Statistical Inference in Political Science
- Course Outline

Required Reading

ALZ Chapter 2

“Islam and Authoritarianism,” by M. Steven Fish, *World Politics*, 2002.

Suggested Reading

ALZ Chapter 1

GH Chapter 1

Wooldridge, Chapter 1

Fox97, Chapter 1

Fox02, Preface, Sections 1.1 and 1.3

Optional Reading

Agresti and Finlay, Chapters 1 and 2

Freedman, et al., Chapters 1 and 2

2 Some Probability Useful for Statistics - Sept. 14

Topics Covered

- Elementary Probability Theory
- Random Variables and Functions of Random Variables

Required Reading

ALZ Chapters 3 and 4

GH Section 2.1

“Racial Prejudice and Attitudes Toward Affirmative Action,” by James H. Kuklinski, Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers, *The American Journal of Political Science*, 1997.

Suggested Reading

Wooldridge Appendix B

Gill, Chapters 7 and 8 (with the exception of 8.3.5, 8.3.8, 8.9)

Fox97 Appendix D.1-D.4

Optional Reading

Gonick and Smith, Chapters 2-5

Freedman et al., Chapters 13-15

Wonnacott & Wonnacott, Sections 3.1-3.6, 4.1-4.5, 5.1-5.4, 6.1-6.3

3 Univariate Statistical Inference (Part 1) - Sept. 21

Topics Covered

- Sampling Distributions
- Point Estimation

Required Reading

ALZ Chapters 6 and 7

GH Sections 2.2, 7.1

Suggested Reading

Wooldridge Appendix C

Agresti and Finlay, Sections 4.3-5.3, 6.1 - 6.5

Fox97 Appendix D.5-D.5.2

Optional Reading

Gonick and Smith, Chapters 6-8

Freedman et al., Chapters 16-26

4 Univariate Statistical Inference (Part 2) - Sept. 28

Topics Covered

- Interval Estimation
- Hypothesis Testing

Required Reading

ALZ Chapter 8

Suggested Reading

Agresti and Finlay, Sections 4.3-5.3, 6.1 - 6.5

Fox97 Appendix D.5-D.5.2

Optional Reading

Gonick and Smith, Chapters 6-8

Freedman et al., Chapters 16-26

5 What is Regression Analysis? - Oct. 5

Topics Covered

- Summarizing and Plotting Bivariate Data
- Review of Joint and Conditional Distributions
- Review of Conditional Expectation
- Non-parametric Regression
- Bias-Variance Tradeoff
- Three Uses for Linear Regression

Required Reading

ALZ Chapter 5 and Sections 9.1, 9.2

Pages 261-272 of, "Measuring Political Preferences," by Lee Epstein and Carol Mershon, *American Journal of Political Science*, 1996.

"Momentous Sprint at the 2156 Olympics," by Andrew J. Tatem, Carlos A. Guerra, Peter M. Atkinson, and Simon I. Hay, *Nature* 2004. (available at course website)

Suggested Reading

Cleveland, Chapter 2 pp. 16 - 33 and Chapter 3 pp. 86-101

Optional Reading

Tufte, Edward. 2001. *The Visual Display of Quantitative Information, 2nd Edition*. Cheshire, CN: Graphics Press.

Cleveland, William S. and Robert McGill. 1987. "Graphical Perception: The Visual Decoding of Quantitative Information on Graphical Displays of Data." (with discussion) *JRSS A*. 150: 192-229.

6 Simple Linear Regression in Scalar Form - Oct. 12

Topics Covered

- Simple Correlation and Least Squares Regression
- Simple Dummy Variable Regression
- Derivation and Formulas
- Properties of Point and Interval Estimators
- Hypothesis Tests

Required Reading

ALZ Chapters 9 and 10

GH Chapters 3 and 4 (but especially Sections 3.1,3.4,3.6,4.1,4.3-4.5), Section 7.1

Suggested Reading

Wooldridge, Chapter 2

Fox97, Sections 4-4.4, 5.1 and 6.1

Fox02, Section 3.4, 4.1.1

Optional Reading

Freedman et al., Chapters 8-12

Wonnacott & Wonnacott, Chapters 11 and 12

Stigler, Steven M. 1986. *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Harvard University Press. Chapters 1,8

7 Linear Regression with Two Explanatory Variables - Oct. 19

Topics Covered

- Dummy Variables and Continuous Regressors
- Interaction Terms
- Correlation and Least Squares Regression
- Added Variable Plots
- Formulas and Intuition

Required Reading

ALZ Chapter 11 and Sections 12.1 and 12.2

GH Chapters 3 and 4, Section 7.2

Review "Islam and Authoritarianism," by M. Steven Fish, *World Politics*, 2002.

Suggested Reading

Wooldridge, Sections 3.1,3.2, 4.1-4.3, 7.1-7.4

Fox97 Sections 5.2, 5.3, 7-7.3.3

Fox02, Section 4.1, 4.2

8 Causality with Least Squares Regression - Oct. 26

Topics Covered

- Linear Regression Causal Model
- Neyman-Rubin Causal Model
- Identification of Causal Effects

Required Reading

Gelman/Hill Chapter 9

“Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment,” by Alan S. Gerber, Donald P. Green, and Christopher W. Larimer. *The American Political Science Review*, 2008.

“Female Socialization: How Daughters Affect Their Legislator Fathers’ Voting on Women’s Issues,” by Ebonya L. Washington, *The American Economic Review*, 2008.

Suggested Reading

Wooldridge 3.3

Fox97 6.2-6.5

Optional Reading

Holland, Paul. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81: 945-960.

Glymour, Clark. 1986. “Comment on Statistics and Causal Inference: Statistics and Metaphysics.” *Journal of the American Statistical Association* 81: 964-966.

Pearl, Judea. 2000. “Epilogue: The Art and Science of Cause and Effect.” in *Causality*. Cambridge University Press.

9 Multiple Regression in Matrix Form - Nov. 2

Topics Covered

- Matrix Algebra
- Vector Calculus
- Derivation of the Least Squares Estimator

Required Reading

Wooldridge Appendix D

Wooldridge Appendix E

Suggested Reading

Gill Chapters 3, 4

Fox97, Appendices C.1, C.3, Sections 9.1 pp. 204-206, 9.2

Optional Reading

Freedman et al., Chapters 26-29

Simon and Blume. 1994. *Mathematics for Economists*. Sections 8.1-8.4

Cleveland. Chapter 4

10 Statistical Inference for Least Squares Regression - Nov. 9

Topics Covered

- Properties of the Least Squares Estimator
- Gauss-Markov Theorem
- Testing and Intervals for Linear Combinations of the Parameters
- Confidence Intervals and Ellipses

Required Reading

ALZ Chapter 12

Wooldridge, Sections 4.4-4.6

Fox97, Sections 9.4

Suggested Reading

Fox02, Sections 4.5 and 4.6

Optional Reading

Kennedy, Appendix A, 418-422

Freedman et al., Chapters 26-29

11 Diagnosing and Fixing Problems I - Nov. 16

Topics Covered

- Nonconstant Error Variance
- Weighted Least Squares
- Heteroskedasticity-Robust Standard Errors
- Correlated Errors
- Generalized Least Squares

Required Reading

ALZ Chapter 14

“The Classical Liberals Were Right: Democracy, Interdependence, and Conflict 1950-1985,” by John R. Oneal and Bruce M. Russett, *International Studies Quarterly*, 1997. 596-627.

Suggested Reading

Fox02, Sections 6.3

Fox97, 12.2, 14.1

Fox02, Web Appendix: “Time-Series Regression and Generalized Least Squares”

Beck, Nathaniel and Katz, Jonathan. 1995. “What to do (and not to do) with Time Series Cross-Sectional Data.” *APSR*. 89: 634-647. (available via JSTOR)

Optional Reading

Berk, Kenneth. 1998. “Regression Diagnostic Plots in 3-D” *Technometrics*. 40: 39-47.

Weisberg, Sanford. 2005. *Applied Linear Regression*. 3rd Edition. Hoboken, NJ: John Wiley.

Wooldridge, Chapters 10-12

12 Diagnosing and Fixing Problems II - Nov. 23

Topics Covered

- Leverage Points
- Outliers
- Influence Points
- Nonnormality
- Nonlinearity

Required Reading

ALZ Chapter 16

Beck, Nathaniel and Simon Jackman. 1998. "Beyond Linearity by Default: Generalized Additive Models." *AJPS*. 42:

Suggested Reading

Fox97, Chapter 11, 12.1,12.3,12.4, 12.7

Fox02, Section 6.1, 6.2, 6.3, and 6.4

Optional Reading

Jackman, Robert W. 1987. "The Politics of Economic Growth in the Industrial Democracies, 1974-80: Leftist Strength or North Sea Oil?" *The Journal of Politics*, Vol. 49, No. 1, pp. 242-256. (available via JSTOR)

Wand, Jonathan; Kenneth Shotts; Jasjeet Sekhon; Walter Mebane; Michael Herron; and Henry Brady. 2001. "The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida." *APSR*. 95: 793-810.

Weisberg, Sanford. 2005. *Applied Linear Regression*. 3rd Edition. Hoboken, NJ: John Wiley.

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

Hastie, Trevor and Rob Tibshirani. 1990. *Generalized Additive Models*. London: Chapman & Hall.

13 Preview of Advanced Causal Inference - Nov. 30

Topics Covered

- Instrumental Variables Analysis

Required Reading

ALZ Chapter 13

Suggested Reading

"Identification of causal effects using instrumental variables," by Angrist, J.D. and Imbens, G.W. and Rubin, D.B., *Journal of the American Statistical Association*, 1996.