

PL SC 503: “Multivariate Analysis for Political Research”

Spring 2009

Professor Christopher Zorn
Department of Political Science
Pennsylvania State University
E-mail: zorn@psu.edu
T-Th 4:15-5:30 p.m.
Fenske Lab, Room 144

Course Description

This is the second (full) course in quantitative methods in Penn State’s political science Ph.D. program. The course introduces students to linear regression models for the analysis of quantitative data, and provides a basis of knowledge for more advanced statistical methods. The course assumes basic math literacy, including familiarity with probability theory, properties of estimators, rudimentary calculus, and linear algebra, as well as mastery of the basic statistics taught in PLSC 502. The bulk of the course will focus on regression models for continuous response variables, and will include discussions of the mathematical bases for such models, their estimation and interpretation, model assumptions and techniques for addressing violations of those assumptions, and topics related to model specification and functional forms. We’ll conclude with an introduction to the idea of maximum likelihood, including an overview of generalized linear models (logit, probit, etc.).

Texts

Required:

Fox, John. 2008. *Applied Regression Analysis and Generalized Linear Models, Second Edition*. Thousand Oaks, CA: Sage Publications. (Click on the title to go to the webpage for the book).

Kennedy, Peter. 2003. *A Guide to Econometrics*, 5th. Ed. Cambridge: MIT Press.

Additional readings as necessary, all of which will be available on ANGEL.TM

The Fox text will be the primary text for the course. It’s expensive, but a good reference. Kennedy is a sort of “Cliff’s Notes” to regression models; you’ll be glad you have it at exam time (or in five years, when you’re revising that dissertation chapter for a journal).

Recommended:

Getting Started with Stata, Release 10. 2007. College Station, TX: Stata Press. A good introduction to **Stata** (should you need one).

Fox, John, and Georges Monette. 2002. *An R and S-Plus Companion to Applied Regression.* Thousand Oaks, CA: Sage Publications. A companion to the Fox text, for S-PlusTM/R users.

King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton: Princeton University Press. Read for background.

Nagler, Jonathan. 1996. "Coding Style and Good Computing Practices." *The Political Methodologist* 6(2):2-8. Contains words to live by.

Some "Econometrics" Texts:

Dougherty, Christopher. 2007. *Introduction to Econometrics*, 3rd Ed. New York: Oxford University Press.

Greene, William. 2008. *Econometric Analysis*, 6th Ed. New York: Prentice-Hall.

Gujarati, Damodar. 2003. *Basic Econometrics*, 4th. Ed. New York: McGraw-Hill.

Hill, R. Carter, William E. Griffiths, and Guay C. Lim. 2007. *Principles of Econometrics*, 3rd Ed. New York: Wiley.

Kmenta, Jan. 1997. *Elements of Econometrics*, 2nd Ed. Ann Arbor, MI: University of Michigan Press.

Maddala, G. S. 2001. *Introduction to Econometrics*, 3rd Ed. New York: Wiley.

Stock, James S. and Mark W. Watson. 2006. *Introduction to Econometrics*, 2nd International Edition. New York: Addison-Wesley.

Wooldridge, Jeffrey. 2005. *Introductory Econometrics: A Modern Approach*, 3rd Ed. Mason, OH: South-Western College Publishing.

Most of these are generally similar to Fox, though with more of an "econometric" flavor and less emphasis on visualization, etc.

Other Good Regression Texts:

Chatterjee, Samprit, and Ali S. Hadi. 2006. *Regression Analysis by Example*, 4th Ed. New York: Wiley.

Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken. 2002. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3rd Ed. Lawrence Erlbaum.

Gelman, Andrew, and Jennifer Hill 2006. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. New York: Cambridge University Press.

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. 2006. *Introduction to Linear Regression Analysis*, 4th Ed. New York: Wiley.

Seber, George A.F., and Alan J. Lee. 2003. *Linear Regression Analysis*, 2nd Ed. New York: Wiley.

Weisberg, Sanford. 2005. *Applied Linear Regression*, 3rd Ed. New York: Wiley.

A Few Other Useful References:

Berry, William D. 1992. *Understanding Regression Assumptions*. Thousand Oaks, CA: Sage Publications.

Fox, John. 1992. *Regression Diagnostics*. Thousand Oaks, CA: Sage Publications.

Lewis-Beck, Michael S. 1995. *Data Analysis: An Introduction*. Thousand Oaks, CA: Sage Publications.

Pindyck, Robert S., and Daniel Rubinfeld. 2000. *Econometric Models and Economic Forecasts*, 4th Revised Ed. New York: McGraw-Hill.

Grading

Grading will be based on a total of 1000 points, divided as follows:

- Lab exercises: Ten worth 50 points each.
- A final paper/project, worth 500 points.

Lab exercises will all be able to be completed using either **Stata**TM or R statistical packages, though you're welcome to adopt whatever other software you care to use. Details for the homeworks and the final project will be announced in class.

Some Other Useful Resources

The **Inter-University Consortium for Political and Social Research** (ICPSR), at the University of Michigan, maintains an extensive archive of data in the social and behavioral sciences. Much of it is accessible via their homepage (<http://www.icpsr.umich.edu>).

The **Political Methodology Section** of the American Political Science Association was created to provide APSA members with an interest in political methodology with a forum in which to meet and discuss ideas. The section publishes a quarterly newsletter (*The Political Methodologist*), a quarterly journal on political methodology (*Political Analysis*), conducts a discussion list on topics relating to political methodology, and maintains an extensive electronic archive of papers, accessible via their homepage (at <http://polmeth.wustl.edu>).

The **Stata**TM homepage (<http://www.stata.com>) is a valuable resource for questions about the Stata statistical software. Similarly, for R users, the **Comprehensive R Archive Network** (CRAN) (<http://cran.r-project.org/>) is the place to go for downloads, packages, and documentation.

Obligatory Statement on Academic Dishonesty

The Department of Political Science, along with the College of the Liberal Arts and the University, takes violations of academic dishonesty seriously. Observing basic honesty in one's work, words, ideas, and actions is a principle to which all members of the community are required to subscribe.

All course work by students is to be done on an individual basis unless an instructor clearly states that an alternative is acceptable. Any reference materials used in the preparation of any assignment must be explicitly cited. Students uncertain about proper citation are responsible for checking with their instructor.

In an examination setting, unless the instructor gives explicit prior instructions to the contrary, whether the examination is in-class or take-home, violations of academic integrity shall consist but are not limited to any attempt to receive assistance from written or printed aids, or from any person or papers or electronic devices, or of any attempt to give assistance, whether the one so doing has completed his or her own work or not.

Lying to the instructor or purposely misleading any Penn State administrator shall also constitute a violation of academic integrity.

In cases of any violation of academic integrity it is the policy of the Department of Political Science to follow procedures established by the College of the Liberal Arts. More information on academic integrity and procedures followed for violation can be found [here](#).

Disabilities

The Pennsylvania State University encourages qualified people with disabilities to participate in its programs and activities and is committed to the policy that all people shall have equal access to programs, facilities, and admissions without regard to personal characteristics not related to ability, performance, or qualifications as determined by University policy or by state or federal authorities. If you anticipate needing any type of accommodation in this course or have questions about physical access, please tell the instructor as soon as possible. Reasonable accommodations will be made for all students with disabilities, but it is the student's responsibility to inform the instructor early in the term. Do not wait until just before an exam to decide you want to inform the instructor of a learning disability; any accommodations for disabilities must be arranged well in advance.

Course Schedule

Some Preliminaries

- **January 13:** *Introductory remarks*
No readings assigned (read Fox, pp. 3-10, for background / review).
- **January 15:** *Regression: A Conceptual Overview*
Readings:
 - Fox, pp. 11-17.

Linear Regression: Basics

- **January 20:** *Bivariate Regression: A (Re)Introduction*
Readings:
 - Fox, pp. 77-86.
- **January 22:** *Bivariate Regression: Inference*
Readings:
 - Fox, pp. 100-105
- **January 27:** *Bivariate Regression: Interpretation*
No readings assigned.
- **January 29:** *Bivariate Regression: R^2 and Model Fit*
Readings:
 - Fox, pp. 92-94.

- Lewis-Beck, Michael S. and Andrew Skalaban. 1990. “When to Use R-Squared.” *The Political Methodologist* 3(2):11-12.
- King, Gary. 1990. “When Not to Use R-Squared.” *The Political Methodologist* 3(2):9-11.
- Luskin, Robert C. 1991. “R-Squared Encore.” *The Political Methodologist* 4(1):21-23.

Multivariate Linear Regression

- **February 3:** *Multivariate Regression*

Readings:

- Fox, pp. 187-197.

Homework One due.

- **February 5:** *Multivariate Regression: Hypothesis Testing and Inference*

Readings:

- Fox, pp. 198-207.
- Kennedy, pp. 60-80.

- **February 10:** *Multivariate Regression: Hypothesis Testing and Inference, II*
No readings assigned.

- **February 12:** *Multivariate Regression: Model Fit and Outliers, I*

Readings:

- Fox, pp. 241-260.

Homework Two due.

- **February 17:** No Class – ISA Meeting in New York.

- **February 19:** *Multivariate Regression: Model Fit and Outliers, II*

No readings assigned

Multiple Regression: Assumptions, Violations, and Remedies

- **February 24:** *Multicollinearity*

Readings:

- Fox, pp. 307-313, 320-332.
- Kennedy, pp. 205-217.

Homework Three due.

- **February 26:** *Errors in Variables*

Readings:

- Fox, pp. 112-118.
- Kennedy, pp. 160-163.

- **March 3:** *Specification Error and Random Regressors*

Readings:

- Fox, pp. 110-112; 210-213.
- Kennedy, pp. 107-109.

Homework Four due.

- **March 5:** *Simultaneity and Endogeneity*

Readings:

- Kennedy, pp. 180-191.

- **March 10:** No Class – Spring Break

- **March 12:** No Class – Spring Break

- **March 17:** *Heteroskedasticity, I*

Readings:

- Fox, pp. 272-277, 290-291.

Homework Five due.

- **March 19:** *Heteroskedasticity, II*

Readings:

- Fox, pp. 296-298; 427-429.
- Downs, George and David Rocke. 1979. "Interpreting Heteroskedasticity." *American Journal of Political Science* 23(November):816-828.

- **March 24:** *Autocorrelation*

Readings:

- Fox, pp. 429-450.
- Kennedy, pp. 139-144.

Homework Six due.

- **March 26: No Class: ICPSR Council.**

Practical Regression: Model Specification and Interpretation

- **March 31:** *Dichotomous Covariates*

Readings:

- Fox, pp. 120-131.
- Kennedy, pp. 248-252.

Homework Seven Due.

- **April 2: No Class – MPSA.**

- **April 7:** *Interaction Terms I*

Readings:

- Fox, pp. 131-140.
- Friedrich, Robert J. 1982. "In Defense of Multiplicative Terms in Multiple Regression Equations." *American Journal of Political Science* 26(November):797-833.

Homework Eight due.

- **April 9:** *Interaction Terms II*

Readings:

- Brambor, Thomas, William R. Clark, and Matt Golder. 2006. “Understanding Interaction Models: Improving Empirical Analyses.” *Political Analysis* 14(1):63-82.

- **April 14:** *(Non-)Linearity and Data Transformations.*

Readings:

- Fox, pp. 50-68; 277-286.

Homework Nine due.

- **April 16:** *Nonparametric Regression.*

Readings:

- Fox, pp. 17-25, 476-517 (read quickly).

Maximum Likelihood Estimation and Generalized Linear Models

- **April 21:** *Maximum Likelihood - Introduction.*

Readings:

- Fox, [Appendix D6](#), pp. 92-95.

Homework Ten due.

- **April 23:** **No Class – Instructor away.**

- **April 28:** *Logit and Probit, I.*

Readings:

- Fox, pp. 335-347.

- **April 30:** *Logit and Probit, II.*

- Fox, pp. 347-355.

Homework Ten due.

- **May 4: Final Projects Due.**