

Yale University
Department of Political Science
EMPIRICAL ANALYSIS & RESEARCH METHODS READING LIST
2011

The objectives of the EARM field are threefold: to encourage graduate students to familiarize themselves with the main techniques of quantitative research employed in political science, to stimulate students to think about how they would design and implement empirical research projects that are both tractable and illuminating, and to train students to become intelligent consumers of quantitative work.

Coursework in the field falls into two categories. The first is the basic three-course sequence PLSC 500, PLSC 503, and PLSC 504 which introduce the main ideas in quantitative research methodologies including sampling, structural modeling, estimation, and measurement. The second category of classes are more specialized courses in the department such as PLSC 507, PLSC 508, and PLSC 512 addressing topics such as experimental methods and nonparameteric statistics in greater depth than offered in the general sequence. Other specialized courses are offered in other departments. More generally, the EARM field spans the breadth of Political Science, and students are expected to be familiar with methodological issues in international relations, comparative politics, American politics, and kindred fields.

What should one know in order to pass the EARM exam? (See the reading list on the next page.) The exam is divided into three sections:

1. Statistical reasoning: students will be asked to interpret quantitative results and to solve/address standard statistical problems in econometrics. This section is mandatory for all exam-takers.
2. Research design: students will be asked to propose a research design that would address a given theoretical question; successful exams will consider problems of theoretical scope, internal and external validity, operationalization, measurement, and data analysis.
3. Critical evaluation of research methodology: students will be asked to read and offer a methodological critique of a brief research paper.

EMPIRICAL ANALYSIS AND RESEARCH METHODOLOGY READING LIST

Here is a reading list of methodologically informative essays that should help you prepare for the exam.

I. STATISTICS

You should have a strong command of probability theory, conditional probability and independence, bayes' rule, discrete distributions, continuous distributions, expectations, variance, and covariance. The following books are ordered from the less to the most mathematical.

Wonnacott, Thomas H. and Ronald J. Wonnacott. 1990. *Introductory Statistics for Business and Economics*. John Wiley & Sons Inc.

Morris H. DeGroot and Mark J. Schervish. 2002. *Probability and Statistics*, Third Edition. Reading, MA: Addison-Wesley.

Casella, George, and Roger L. Berger. 2002. *Statistical Inference*. Duxbury, Thomson Learning.

II. REGRESSION

a. Mechanics of least squares and the classical linear model

Wooldridge, Jeffrey M. 2009. *Introductory Econometrics. A Modern Approach*. Chapters 2, 3 and appendix E.

Freedman, David. 2006. *Statistical Models: Theory and Practice*. New York: Cambridge University Press. Chapters 2-4. (Chapters 2-5 in the 2009 revised edition.)

Greene, William. 2002. *Econometric Analysis*, 5th Edition. New York: Prentice-Hall. Chapter 6.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. Chapter 3.

Cameron, Colin A. and Pravin K. Trivedi. 2005. *Microeconometrics, Methods and Applications*. Cambridge University Press. Chapter 4.

b. Statistical Inference and hypothesis testing

Cameron, Colin A. and Trivedi, Pravin K. 2005. *Microeconometrics, Methods and Applications*. Cambridge University Press. Chapter 7.

Greene, William. 2002. *Econometric Analysis*, 5th Edition. New York: Prentice-Hall. Chapter 7.

Gerber, Alan S. and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*. Yale manuscript. Chapter 3.

c. OLS asymptotics

Greene, William. 2002. *Econometric Analysis*, 5th Edition. New York: Prentice-Hall. Chapter 9.

d. Binary choice and interactions

Wooldridge, Jeffrey M. 2009. *Introductory Econometrics. A Modern Approach*. Chapter 7.

Chatterjee, Samprit and Alis S. Hadi. 2006. *Regression analysis by example*. John Wiley & Sons. Chapter 5.

III. MODEL SPECIFICATION

Bartels, L. 1990. "Five Approaches to Model Specification". *The Political Methodologist* 3: (No. 2) 2-6.

LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review*, 76: 604-20.

Leamer, EE. 1983, "Let's Take the Con out of Econometrics". *The American Economic Review*, 73(1): 31-43.

IV. HETEROSKEDASTICITY, MEASUREMENT ERROR AND OTHER DATA PROBLEMS

King, G. 1986. "How Not to Lie With Statistics: Avoiding Common Mistakes in Quantitative Political Science". *American Journal of Political Science*, 30(No.3) 666-687.

Greene, William. 2002. *Econometric Analysis*, 5th Edition. New York: Prentice-Hall. Chapters 11-12.

Given the adverse statistical consequences of measurement error, a large body of statistics discusses the analysis of models with latent variables. For an accessible introduction, see:

Long, J. Scott. 1983 *Covariance Structure Models: An Introduction to LISREL*. Beverly Hills, Calif.: Sage Publications

Among the many applications, see:

Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder Jr. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting". *American Political Science Review* 102:215-232.

V. TIME SERIES

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. 2nd edition, Cambridge, MA: MIT Press

Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press. Chapter 5

Hamilton, James D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

VI. PANEL DATA

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. Cambridge, MA: MIT Press. (Section 6.3.1, Chapter 10)

Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press. Chapter 5.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics* (February).

VII. BAYES

Some familiarity with the Bayesian perspective on statistics is useful. Among the great many texts that are out there, one of the most accessible is:

Iverson, G.R. 1984. *Bayesian Statistical Inference*. Beverly Hills: Sage Publications

Bolstad, William. 2007. *Introduction to Bayesian Statistics*. John Wiley & Sons.

VIII. CORE CONCEPTS IN RESEARCH DESIGN

Gerber, Alan S., Donald P. Green, and Edward H. Kaplan. 2004. *The Illusion of Learning from Observational Research*. In Ian Shapiro, Rogers Smith, and Tarek

Massoud, eds., *Problems and Methods in the Study of Politics*. New York: Cambridge University Press, pp. 251-73.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge, UK: Cambridge University Press. Chapters 1-2.

Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press. Chapters 1-3.

Gerber, Alan S. and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*. Yale manuscript. Chapters 1 and 2.

Campbell, D.T. and D.W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix". *Psychological Bulletin*, 56: 81-105

Cook, T.D. and Campbell, D.T. 1979. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Chicago: Rand McNally College Pub. Co.

Lieberson, Stanley. 1991. Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases. *Social Forces*. 70: 307-320.

IX. CAUSAL INFERENCE AND EXPERIMENTATION

a. Experimentation in political science

Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia (eds.) *Cambridge Handbook of Experimental Political Science*. Cambridge University Press. Forthcoming.

b. Using covariates in experimental design and analysis

Gerber, Alan S. and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*. Yale manuscript. Chapter 4.

c. Non-compliance

Gerber, Alan S. and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*. Yale manuscript. Chapters 5 and 6.

d. Attrition

Gerber, Alan S. and Donald P. Green, *Field Experiments: Design, Analysis, and Interpretation*. Yale manuscript. Chapter 7

e. Heterogeneous treatment effects

Gerber, Alan S. and Donald P. Green, Field Experiments: Design, Analysis, and Interpretation. Yale manuscript. Chapter 9

f. Mediation

Gerber, Alan S. and Donald P. Green, Field Experiments: Design, Analysis, and Interpretation. Yale manuscript. Chapter 10

X. INSTRUMENTAL VARIABLE STRATEGIES

Angrist, Joshua and Jorn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton, NJ: Princeton University Press. Chapter 4.

Angrist, Joshua, Guido Imbens, and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of American Statistical Association* Vol. 91 No.434 (June):444-455.

Morgan, Stephen L. and Christopher Winship. 2007. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge, UK: Cambridge University Press. Chapter 7.

Wooldridge, Jeffrey. 2010. Econometric Analysis of Cross-Section and Panel Data. 2nd edition. Cambridge: MIT Press, pp. 621-636, Chapter 5.

Sovey, Allison and Donald Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* Vol. 55, No. 1.Pp. 188-200.

Conley, Tim, Christian Hansen, and Peter E. Rossi. 2007. "Plausibly Exogenous." *Review of Economics and Statistics*, in print.

Altonji, Joseph, Todd E. Elder, and Christopher Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* Vol. 113:151-184.

XI. REGRESSION DISCONTINUITY DESIGN AND MATCHING

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.

Angrist, Joshua and Jorn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An

Empiricist's Companion. Princeton, NJ: Princeton University Press. Chapter 6.

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

Imbens, Guido and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142:615-635.

Green, Donald P., Terence Y. Leong, Holger L. Kern, Alan S. Gerber, and Christopher W. Larimer. 2009. Testing the Accuracy of Regression Discontinuity Analysis Using Experimental Benchmarks. *Political Analysis* 17(4): 400-417.

Morgan, Stephen L. and Christopher Winship. 2007. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge, UK: Cambridge University Press. Chapters 4-5.

Sekhon, Jasjeet S. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." *Annual Review of Political Science* 12: 487-508.

Sekhon, Jasjeet. 2008. The Neyman-Rubin Model of Causal Inference and Estimation via Matching Methods. *The Oxford Handbook of Political Methodology* (Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds.) pp.271-299.

Wooldridge, Jeffrey. 2010. *Econometric Analysis of Cross-Section and Panel Data*. 2nd edition. Cambridge: MIT Press, pp. 614-621.

XII. LIMITED DEPENDENT VARIABLES AND MAXIMUM LIKELIHOOD ESTIMATION

King, Gary. 1998. *Unifying Political Methodology*. Ann Arbor: University of Michigan Press. Chapters 2-4.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks, CA: Sage Publications. Chapters 1-2.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. Cambridge, MA: MIT Press. Chapter 13.

XIII. NORMAL REGRESSION IN MLE AND EXTENSIONS

(Probit, Heteroskedastic Regression, Ordered Probit, Censored Regression, Truncated Regression, Selection Models)

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks, CA: Sage Publications. Chapters 3-5.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. Cambridge, MA: MIT Press. Chapters 15-17.

XIV. PROBABILISTIC CHOICE MODELS

(Logit, Multinomial Logit, Conditional Logit, Nested Logit, and Multinomial Probit)

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks, CA: Sage Publications. Chapter 6.

XV. EVENT COUNT MODELS

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*, Thousand Oaks, CA: Sage Publications. Chapter 8.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. Cambridge, MA: MIT Press. Chapter 19.

XVI. DURATION MODELS

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd edition. Cambridge, MA: MIT Press. Chapter 20.

The Oxford Handbook of Political Methodology covers other topics of interest, such as ecological inference, survival models, multilevel models, etc.