Causal Inference: Methods for Program Evaluation and Policy Research
APSTA-GE 2012.1.001
Fall 2015

Course Time and Location:
     Friday 9:30-12:15
     Meyer 122
Instructor: Jennifer Hill
          804 Kimball Hall (246 Greene St)
Office Hours: Wednesday 3:00-4:00

Course Description and Prerequisites:
This goal of this course is to provide students with a basic knowledge of how to perform some more advanced statistical methods useful in answering policy questions using observational or experimental data. It will also allow them to more critically review research published that claims to answer causal policy questions. The primary focus is the challenge of answering causal questions, that is, those that take the form "Did A cause B?" using data that do not conform to a perfectly controlled randomized study. Examples from real public policy studies will be used throughout the course to illustrate key ideas and methods. First, we will explore how best to design a study to answer causal questions given the logistical and ethical constraints that exist. We then discuss several approaches to drawing causal inferences from observational studies including propensity score methods, instrumental variables, regression discontinuity, difference in differences, and fixed effects models. If time allows we will also discuss some simple forms of sensitivity analysis.

Pre-requisites:
The prerequisite for the course is two semesters of quantitative methods at the level of RESCH-GE.2003, RESCH-GE.2004 (or the equivalent as approved by the instructor). Students should understand the properties of regression including use of indicator (dummy) variables, interactions, and transformation. Exposure to logistic regression is also required. Experience with Stata is extremely helpful though not formally required.

Grading:
Grading will be based on approximately 7 graded assignments (totaling 70%) and one final project or final exam (30%) that will involve both data analysis and a thoughtful description of both the analysis and the findings. A description of the final exam and final project as well as detailed guidelines for the project will be provided in a document posted on the Classes website.

Class meeting times, organization, deadlines
Each class is scheduled to run for 2 hours and 45 minutes. Technically, we are required to meet only 2 hours each week. 10-15 minutes of the extra time will go towards making up for a class that we will have to miss in December (see below). The rest of the time will be used for additional Stata instruction, help with homework, or more general questions/office hours as the need arises. Therefore you should plan on being in the classroom each week until 11:45. Typically, the remaining time (from 11:45-12:15) will be optional.

The class will meet every Friday that the university is open during the semester (that is, all but the day after Thanksgiving) with the exception of Friday October 9th. The last day of class will be Friday December 11th. The final project/final exam will be due on Friday December 18th at noon. *A hardcopy of your final must be handed in to my office or mailbox; electronic copies will not be accepted.*

Reading materials
There is no required textbook for the course:

The required readings are either available through e-journals through the library or will be posted on the Classes website. In addition, however, the following books are recommended:

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

Guo, Shenyang and Fraser, Mark W. (2010) *Propensity Score Analysis*, Sage Publications.

Morgan, S. and Winship, C. (2007) *Counterfactuals and Causal Inference: Methods and Principles for Social Research,* Cambridge University Press

Angrist and Pischke (2009) *Mostly Harmless Econometrics*, Princeton University Press.

Imbens and Rubin (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences,* Cambridge University Press.

Outline of course Lessons and readings:
The following outline describes the Lessons that will be covered along with anticipated associated readings. It corresponds roughly to the course weeks though we may end up adjusting time spent on each Lesson as we go. Readings highlighted with an * are recommended, not required. All readings not available on the web will be put on electronic reserve through Bobst library (BL). Web addresses for the others are provided.

1) **Motivation: What's all the observational vs. randomized fuss about?**
   Simple randomized experiments (theory and practice) and the Rubin Causal Model

*Winship & Morgan, Chapter 2 (posted on Classes under Lesson 1)

*Leamer, Edward (1983) "Let's take the con out of econometrics", *American Economic Review*, 73(1): 31-43 [Available at www.jstor.org]

*Hill, J., Reiter, J., and Zanutto, E. (2004) "A comparison of experimental and observational data analyses" *Applied Bayesian Modeling and Causal Inference from an Incomplete-Data Perspective.* Edited by Andrew Gelman and Xiao-Li Meng. West Sussex, England: Wiley.
[Available on Classes under Lesson 1]

*Holland, Paul W. (1986), ``Statistics and causal inference (with discussion)'', *Journal of the American Statistical Association*, 81: 945-970 [Available at www.jstor.org]

*Gelman, Andrew and Hill, Jennifer (2007) Chapters 3 and 4, of *Data Analysis Using Regression and Multilevel/Hierarchical Models,* Cambridge University Press

2) Randomized experiments (including Randomized Block and Matched Pairs Designs) and complications that make them look like observational studies

Gelman, A., Hill, J. and Reiter, J. (2013) Draft Chapter "Causal Inference Basics and Randomized Experiments" for *Regression and Other Stories*, to be published by Cambridge University Press in 2014
available as gelman.hill.reiter.draft.RandomizedExperiments.pdf on Classes under Lesson 2

Katz, L.F., Kling, J.R., and Liebman, J.B. (2001) "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment" *The Quarterly Journal of Economics* 116: 607-654.
available on Classes under Lesson 2

Guo and Fraser, Chapter 2, pp 21-36
available on Classes under Lesson 2

*Gelman, Andrew and Hill, Jennifer (2007) Chapter 9, Sections 9.1-9.4, of *Data Analysis Using Regression and Multilevel/Hierarchical Models,* Cambridge University Press

*Imbens, G., Rubin, D., and Sacerdote, B. (2001) "Estimating the Effect of Unearned Income on Labor Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players" *American Economic Review* 91(4): 778 (available through e-journals)

*Rubin, D. (1990) "Formal modes of statistical inference for causal effects" *Journal of Statistical Planning and Inference* 25: 279-292. (available through e-journals)

*Sobel, Michael E. (1996), "An introduction to causal inference", *Sociological Methods and Research*, Vol. 24, Iss. 3; p. 353-379 (LL) (available through e-journals)

*Rosenbaum, P. (2002) *Observational Studies*, *2nd ed.*, New York: Springer, Chapter 2

3) Observational Studies and simple ways of adjusting for covariates

Gelman, A., Hill, J. and Reiter, J. (2013) Draft Chapter "Causal Inference with Observational Studies Assuming All Confounders Have Been Measured" for *Regression and Other Stories*, to be published by Cambridge University Press in 2014 (read until section on propensity scores)
available as gelman.hill.reiter.draft.ObservationalStudies1.pdf on Classes

Rosenbaum, P. (2002) *Observational Studies*, *2nd ed.*, New York: Springer, Chapter 1
 [available on Classes under Lesson 3]

Donohue, J. J., and S.D. Levitt (2001) "The impact of legalized abortion on crime" *The Quarterly Journal of Economics,* 116(2): 379-420
[Available through e-journals]

*Gelman, Andrew and Hill, Jennifer (2007) Chapters 3&4 of *Data Analysis Using Regression and Multilevel/Hierarchical Models* in press at Cambridge University Press

*Winship, Christopher and Michael Sobel (2004) "Causal Inference in Sociological Studies" in Handbook of Data Analysis edited by Melissa Hardy and Alan Bryman, London: Sage Publications, 36-38

*Rubin, D. (1977) "Assignment to Treatment Groups on the Basis of a Covariate" *Journal of Educational Statistics*, 2: 1-26

*LaLonde, R. (1986) Evaluating the Econometric Evaluations of Training Programs, *American Economic Review*, 76: 604-620 [Available at www.jstor.org]

*Rosenbaum, P. (2002) "Covariance adjustment in randomized experiments and observational studies", *Statistical Science,* 17(3): 286-327

4) Propensity Score Approaches – Introduction

Gelman, A., Hill, J. and Reiter, J. (2014) Draft Chapter "Causal Inference with Observational Studies Assuming All Confounders Have Been Measured" for *Regression and Other Stories*, to be published by Cambridge University Press in 2014 (section on propensity score matching) available as gelman.hill.reiter.draft.ObservationalStudies2.pdf on Classes

Bingenheimer, J.B., Brennan, R.T., and Earls, F.J. (2005) "Firearm violence exposure and serious violent behavior," *Science* 308: 1323-1326 [Available on Classes under Lesson 4]

Rosenbaum, PR and D B. Rubin (1985) "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score", *The American Statistician*, 39: 33-38 [Available at www.jstor.org]

Schafer, J. and Kang, J. (2008) "Average Causal Effects from Nonrandomized Studies: A Practical Guide and Simulated Example" *Psychological Methods* [Available on Classes under Lesson 4]

*Rosenbaum, Paul R. and Rubin, Donald B. (1984) "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score" *Journal of the American Statistical Association*, 79: 516—524 [Available at www.jstor.org]

5) Propensity Score Approaches – Practice (Stata)

Hill, J. (2008) "Discussion of research using propensity-score matching: Comments on 'A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003' by Peter Austin," *Statistics in Medicine*, 27: 2055-2061.

Guo and Fraser, Chapter 5, pp 127-149, 154-158, 161-162, 167-186
Chapter 5 available on Classes under Lesson 4

*D'Agostino, R (1998) "Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group" *Statistics in Medicine*, 17: 2265-2281.

*O'Keefe, Suzanne (2004) "Job creation in California's enterprise zones: a comparison using a propensity score matching model" Journal of Urban Economics, 55: 131-150.
[available on Classes]

*Foster, M. (2003) "Propensity Score Matching: An Illustrative Analysis of Dose Response"
Medical Care 41: 1181-1192
[Available through e-journals]

*Dehejia, Rajeev H. and Wahba, Sadek (1999) "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of the American Statistical Association*, 94: 1053—1062 [Available at www.jstor.org]

6) Instrumental Variables Models – Introduction and Theory

Gelman, A., Hill, J. and Reiter, J. (2013) Draft Chapter "Observational Studies that Relax or Extend the Ignorability Assumption" for *Regression and Other Stories*, to be published by Cambridge University Press in 2014, section on instrumental variables
available as gelman.hill.reiter.draft.ObservationalStudies2.pdf on Classes

Gennetian, L., Bos, J., Morris, P. (2002) "Using Instrumental Variables Analysis to Learn More from Social Policy Experiments" Manpower Demonstration Research Corp. working paper
[available on Classes]

Angrist, J D., Imbens, G W. and D B. Rubin, (1996) "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91: 444-472 [Available at www.jstor.org]

*Angrist, J D. (1990) "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review,* 80: 313-336 [Available at www.jstor.org]

7) Instrumental Variables Models – Practice (Stata)

Angrist JD, Evans WN (1998), "Children and their parents' labor supply: Evidence from exogenous variation in family size", *American Economic Review* 88(3): 450-77
[available at www.jstor.org]

*E. Michael Foster. (2000) "Is more better than less? An analysis of children's mental health services" *Health Services Research.* Chicago: Vol. 35, Iss. 5; p. 1135

*Levitt, Steven D. 1996. "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation." *Quarterly Journal of Economics*, 111(2): 319-51. [Available at www.jstor.org]

8) Regression Discontinuity

Gelman, A., Hill, J. and Reiter, J. (2013) Draft Chapter "Causal Inference with Observational Studies Assuming All Confounders Have Been Measured" for *Regression and Other Stories*, to be published by Cambridge University Press in 2014 (section on regression discontinuous) available as gelman.hill.reiter.draft.ObservationalStudies1.pdf on Classes

Brian A. Jacob, Lars Lefgren (2004) "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis" *Review of Economics and Statistics* 86(1)

Shadish, Cook & Campbell (2002) "Regression Discontinuity Designs" Chapter 7 in *Experimental and Quasi-Experimental Designs* Boston: Houghton Mifflin Co. (posted on Classes with lecture)

 9) Difference in Differences/ Fixed Effects models

Gelman, A., Hill, J. and Reiter, J. (2013) Draft Chapter "Causal Inference with Observational Studies Assuming All Confounders Have Been Measured" for *Regression and Other Stories*, to be published by Cambridge University Press in 2014 (section on DID and FE) available as gelman.hill.reiter.draft.ObservationalStudies1.pdf on Classes

Chapter 18 on Panel Data Model in Ashenfelter book *Statistics and Econometrics* (published by Wiley, 2003), pp. 262-273 (available on Classes under Lesson 6)

Angrist, J. D., and Krueger, A. (1999), "Empirical Strategies in Labor Economics," in Orley Ashenfelter and David Card (eds), *Handbook of Labor Economics*, Vol. 3A, Amsterdam: Horth-Holland (available on Classes under Lesson 6) **pp 16-24**

Bogart & Cromwell. "How much is a neighborhood school worth?" *J. Urban Economics* 47 [Available through e-journals]

Aaronson, Daniel. (1998) "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes," *The Journal of Human Resources*, 33(4): 915-946 [Available through e-journals]

*Card, D. and A. Krueger (1994) "Minimum Wages and Employment: A Case Study of the Fast-food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84(4): 772-784. [Available at www.jstor.org]

*Meyer, B. (1995) "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, 13(2): 151-161 [Available through e-journals]

*Korenman and Neumark (1991) "Does Marriage Really Make Men More Productive?" Journal of Human Resources, 26(2): 282-307
[available through e-journals]