

2014 Advanced Causal Inference Workshop Detailed Schedule with Readings

Advanced Workshop Faculty

Donald B. Rubin (Harvard University, Department of Statistics)

Donald Rubin is John L. Loeb Professor of Statistics, Harvard University. His work on the “Rubin Causal Model” is central to modern understanding of when one can and cannot infer causation from regression. Principal research interests: statistical methods for causal inference; Bayesian statistics; analysis of incomplete data. Web page, with link to CV: www.stat.harvard.edu/faculty_page.php?page=rubin.html; Wikipedia: http://en.wikipedia.org/wiki/Donald_Rubin

Justin McCrary (University of California, Berkeley, Law School)

Justin McCrary is Professor of Law, University of California, Berkeley. Principal research interests: crime and urban problems, law and economics, corporations, employment discrimination, and empirical legal studies. Web page with link to CV: <http://www.econ.berkeley.edu/~jmccrary/>.

Jonathan N. Katz (California Institute of Technology)

Jonathan Katz is Kay Sugahara Professor of Social Sciences and Statistics at Caltech. Co-editor: *Political Analysis*. Principal research interests: American politics, political methodology; formal political theory. Web page with link to CV: <http://jkatz.caltech.edu/>.

Advanced Workshop Outline

Wednesday August 13 (Don Rubin)

Choosing estimands (the science). Implications of choice of estimand for choice of method. Principal stratification. Flexible matching methods. Multiple imputation of missing potential outcomes. And whatever else Don thinks he should cover, in the allotted time.

General Reading: Imbens and Rubin, *Causal Inference in Statistics and Social Sciences* (draft 2014), chapters 1-8 (chapter 2 is background and can be skipped).

On estimands: Rubin, Donald (2005), Causal Inference Using Potential Outcomes: Design, Modeling, Decisions, 100 *Journal of the American Statistical Association* 322-331.

On principal stratification:

Comment: Principal stratification was first used (although the term was developed later) for instrumental variable estimates of “local average treatment effects” for treatment with non-compliance. See Angrist, Joshua, Guido Imbens, and Donald Rubin (1996), 91 *Journal of the American Statistical Association* 444-455. The groups relevant for this “causal IV” analysis (always takers, never takers, compliers, and defiers) can be seen as principal strata. If you are not familiar with this use of principal stratification, Imbens and Rubin chs. 24-25 will be useful background.

Meulli, Fabrizia, and Donald Rubin (2003), Assumptions Allowing the Estimation of Direct Causal Effects, 112 *Journal of Econometrics* 79-87. [**Note:** we have not posted the long, complex paper by Peter Adams et al. that Meulli and Rubin are commenting on; you need to know only that they estimate something they call a “direct” causal effect of wealth on mortality by conditioning on health.]

Frumento, Paolo, Fabrizia Mealli, Barbara Pacini, and Donald Rubin (2012), Evaluating the Effect of Training on Wages in the Presence of Noncompliance, Nonemployment, and Missing Outcome Data, 107 *Journal of the American Statistical Association* 450-466.

On other topics: Rubin, Donald B. (2008), For Objective Causal Inference, Design Trumps Analysis, 2 *Annals of Applied Statistics* 808-840.

Thursday August 14 (Justin McCrary)

Conducting simulation studies. Inference and testing using the bootstrap, including adapting bootstrap methods to your research design. Choosing among balancing methods: Matching, reweighting, and regression adjustment. Topics in regression discontinuity design: nonparametric estimation; Local linear regression and density estimation; choosing bandwidth and assessing sensitivity to bandwidth choice.

Readings:

Simulation: A. Colin Cameron and Pravin Trivedi (2010), *Microeconometrics Using Stata*, chapter 13

Bootstrap: Cameron and Trivedi (2010), *Microeconometrics Using Stata*, chapter 13

Russell Davidson and James G. MacKinnon (2005), Bootstrap Methods in Econometrics, in Kerry Patterson and Terence Mills, eds., Palgrave Handbook of Econometrics, vol. 1, chapter 25.

alternative to Davidson and MacKinnon: Bruce Hansen, *Econometrics* (2013), chapters 10

Nonparametric Regression: Hansen (2013), *Econometrics* chapters 11, 12

Regression Discontinuity: John DiNardo and David Lee (2010), "Program Evaluation and Research Designs", in *Handbook of Labor Economics*

Friday August 14 (Jonathan Katz)

Topics in causal inference with panel data, including time-series-cross-sectional (TSCS) data. Topics will include issues of unit heterogeneity, specification of dynamics, synthetic matching, and marginal structural models, and which standard errors to use.

Lunch talk: Advice from a journal editor on what to do (and not do) (Jonathan Katz is co-editor in chief of *Political Analysis*).

Readings:

- Joshua D. Angrist, and Jorn-Steffen Pischke (2009),. *Mostly Harmless Econometrics: An Empiricist's Companion*, chapter. 5
- Nathaniel Beck and Jonathan Katz (1995). What To Do (and Not To Do) with Times-Series–Cross-Section Data in Comparative Politics, 89 *American Political Science Review* 634–647.
- Nathaniel Beck and Jonathan Katz (2011), Modeling Dynamics in Time-Series–Cross-Section Political Economy Data. 14 *Annual Review of Political Science* 331-352.
- Nathaniel Beck, Jonathan N. Katz, and Richard Tucker (1998): Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable, 42 *American Journal of Political Science* 1260-1288.
- Andrew Gelman and Jenifer Hill (2007), *Data Analysis using Regression and Multilevel/Hierarchical Models* Chapters: 11-13
- Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan (2004), How Much Should We Trust Differences in Differences Estimates? 119 *Quarterly Journal of Economics* 249-275.

Additional recommended readings:

- R. Michael Alvarez, Delia Bailey, and Jonathan N. Katz. 2008. "The Effect of Voter Identification Laws on Turnout." Caltech Social Science Working Paper, Number 1267R.
- Nathaniel Beck and Jonathan N. Katz. 2007. "Random Coefficient Models for Time-Series-Cross-Section Data: Monte Carlo Experiments." 15 *Political Analysis* 182–195.
- Cheng Hsiao. 2003. *Analysis of Panel Data*. 2nd Edition. Chapters 1-4.
- Symposium on Research Design and Method in International Relations (2001). 55(2) *International Organization*: Donald P. Green, Soo Yeon Kim, and David H. Yoon: Dirty Pool. pp. 441-468 , and Nathaniel Beck and Jonathan N. Katz: Throwing out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon. pp. 487-495

Workshop Organizers

Bernard Black (Northwestern University, Law and Kellogg School of Management)

Bernie Black is Nicholas J. Chabraja Professor at Northwestern University, with positions in the Law School and Kellogg School of Management. Principal research interests: law and finance, international corporate governance, health law and policy; empirical legal studies. Papers on SSRN: <http://ssrn.com/author=16042>.

Mathew McCubbins (Duke University)

Professor of Political Science and Law at Duke University, with positions in the Law School and the Political Science Department, and director of the Center for Law and Democracy. Principal research interests: democratic institutions, legislative organization; behavioral experiments, communication, learning and decisionmaking; statutory interpretation, administrative procedure, research design; network economics. Web page with link to CV: www.mccubbins.us. Papers on SSRN: <http://ssrn.com/author=17402>.