

Econometrics, Harmless and Otherwise

Alberto Abadie
MIT and NBER

Josh Angrist
MIT and NBER

Chris Walters
UC Berkeley

ASSA January 2020

We discuss econometric techniques for cross-section causal inference, covering recent developments, old favorites, and many empirical applications. Our focus is on conceptual matters and research design rather than statistical programming. Topics include randomized trials, matching, instrumental variables, differences-in-differences, synthetic controls, regression discontinuity designs, structural models for policy evaluation, and Bayesian and machine learning methods. The course consists of nine 80 minutes sessions, two on the first day, five on the second, two on the third.

Lecture 1: Causality, Potential Outcomes, and Treatment Effects in Randomized Studies (Abadie)

Causality, counterfactuals and potential outcomes
Randomized experiments, Fisher's exact test
Threats to internal and external validity in randomized experiments

Lecture 2: Making Regression Make Sense (Angrist)

Causal control
The long and short of regression anatomy
Omitted Variables Bias
Hey, where d'ya go to school?

Lectures 3: Old School IV (Angrist)

IV and OVB
Two-stage least squares
Grouped data and two-sample IV
Beating the bias of 2SLS
Welcome to the machine

Lecture 4: Mastering Modern IV (Angrist)

IV with heterogeneous potential outcomes
IV in RCTs
External validity
All my children

Lecture 5: Matching and the Propensity Score (Abadie)

Selection on observables
Matching, subclassification, and propensity score methods
Comparison of estimators using the National Supported Work Demonstration data
What to match on: a brief introduction to Directed Acyclic Graphs (DAGs)

Lecture 6: Differences-in-Differences and Synthetic Controls (Abadie)

Motivation: difference-in-differences estimation in comparative case studies
Estimation and inference with synthetic controls
Applications

Lecture 7: Regression Discontinuity Designs (Walters)

Theory

- Sharp RD identification
- RD diagnostics
- RD estimation: global and local
- Fuzzy RD
- Regression kink designs

Applications

- The incumbency advantage
- Sheepskin effects: signaling vs. human capital

Lecture 8: Selection Models and Policy Evaluation (Walters)

Theory

- Selection model basics
- Control function estimation
- Connections to IV: LATE equivalence and extrapolation
- Marginal treatment effects and policy evaluation

Application

- Head Start counterfactuals

Lecture 9: Bayesian and Machine Learning Methods (Walters)

Theory

- Empirical Bayes shrinkage
- Bayes vs. empirical Bayes
- Machine learning techniques

Applications

- Teacher and school evaluation
- Social spillovers at the movies

READINGS

Texts: J.D. Angrist and J.S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

J.D. Angrist and J.S. Pischke, *Mastering 'Metrics: The Path from Cause to Effect*, Princeton University Press, 2014.

Overview article: Abadie, A. and M. Cattaneo (2018), "Econometrics Methods for Program Evaluation," *Annual Review of Economics* vol. 10, 465–503.

Many of the readings are from MHE and MM. Journal articles are in JSTOR. Working papers are available from online sources.

CAUSALITY, POTENTIAL OUTCOMES, AND THE ESTIMATION OF TREATMENT EFFECTS IN RANDOMIZED STUDIES

MM Chapter 1; MHE Chapter 2

Duflo, E., R. Glennerster and M. Kremer (2008), "Using Randomization in Development Economics Research: A Toolkit," in T.P. Schultz and J.A. Strauss eds. *Handbook of Development Economics*, vol. 4. Elsevier Science.

Imbens, G.W. and J.M. Wooldridge (2009) "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, vol. 47(1), 5-86.

Imbens, G.W. and D.B. Rubin (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. Chapters 1-10

Krueger, A. (1999), "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics*, vol. 114, 497-532.

MAKING REGRESSION MAKE SENSE

MM Chapter 2; MHE Chapters 1-2 and 3.1-3.2.

S.B. Dale and A.B. Krueger, "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *The Quarterly Journal of Economics* 117, November 2002, 1491-152

OLD SCHOOL INSTRUMENTAL VARIABLES

MM Chapter 3, MHE Section 4.1

J. Angrist and A. Krueger, "Instrumental Variables and the Search for Identification," *Journal of Economic Perspectives*, Fall 2001.

J. Angrist, "Grouped Data Estimation and Testing in Simple Labor Supply Models," *Journal of Econometrics*, February/March 1991.

J. Angrist, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review*, June 1990.

Inoue, Atsushi and G.Solon, "Two-Sample Instrumental Variables Estimators," *The Review of Economics and Statistics*, August 2010.

2SLS Mistakes:

MHE, Section 4.6.1.

The Bias of 2SLS

MHE Section 4.6.4

J. Angrist and A. Krueger, "Does Compulsory Schooling Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics* 106, November 1991, 979-1014.

J. Angrist, G. Imbens, and A. Krueger, "Jackknife Instrumental Variables Estimation," *Journal of Applied Econometrics* 14(1), 57-67.

Flores-Lagunes, Alfonso, "Finite-Sample Evidence on IV Estimators with Weak Instruments," *Journal of Applied Econometrics* 22, 2007, 677-694.

J. Angrist and B. Frandsen, "Machine Labor," NBER Working Paper No. 26584, December 2019.

MASTERING MODERN IV

MM Chapter 3, MHE Section 4.4

G. Imbens and J. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, March 1994.

J. Angrist, G. Imbens, and D. Rubin, "Identification of Causal Effects Using Instrumental Variables," with comments and rejoinder, *JASA*, 1996.

Models with Variable and Continuous Treatment Intensity

MHE Section 4.5.3

J. Angrist and G. Imbens, "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *JASA*, June 1995.

J. Angrist, K. Graddy, and G. Imbens, "The Interpretation of instrumental Variables Estimations in Simultaneous Equations Models with an Application to the Demand for Fish," *Rev. Ec. Studies* 67 (2000), 499-527.

External Validity

J. Angrist, V. Lavy, and Analia Schlosser, "Multiple Experiments for the Causal Link Between the Quantity and Quality of Children," *The Journal of Labor Economics*, October 2010.

J. Angrist and I. Fernandez-Val, "Extrapo-LATEing: External Validity and Overidentification in the LATE Framework," in *Advances in Econometrics Theory and Applications*, Tenth World Congress, Volume III, 2013.

MATCHING AND THE PROPENSITY SCORE

Abadie, A. and G.W. Imbens (2006), "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, vol. 74, 235-267.

Abadie, A. and G.W. Imbens (2011), "Bias Corrected Matching Estimators for Average Treatment Effects," *Journal of Business and Economic Statistics*, vol. 29, 1-11.

Abadie, A. and G.W. Imbens (2012), "A Martingale Representation for Matching Estimators," *Journal of the American Statistical Association*, vol. 107, 833-843.

- Abadie, A. and G.W. Imbens (2016), "Matching on the Estimated Propensity Score," *Econometrica*, vol. 84, 781-807.
- Dehejia, R.H. and S. Wahba (1999), "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, vol. 94, 1053-1062.
- Heckman, J.J., H. Ichimura and P.E. Todd (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, vol. 64, 605-654.
- Hernán, M. and J. Robins (2012), *Causal Inference*. Available online at: <http://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/>
- Hirano, K., G.W. Imbens, and G. Ridder (2003), "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, vol. 71, 1161-1189.
- Imbens, G.W. (2004), "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *Review of Economics and Statistics*, vol. 86(1), 4-29.
- Imbens, G.W. and D.B. Rubin (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. Chapters 12-18.
- Newey, W.K. (1994), "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, vol. 62, 1349-1382.
- Newey, W.K. and D.L. McFadden (1994), "Large Sample Estimation and Hypothesis Testing." In: Engle, R.F., McFadden, D.L. (Eds.), *Handbook of Econometrics*, vol. IV. Amsterdam: Elsevier Science.
- Pearl, J. (2009), *Causality* (second edition). Cambridge University Press.
- Rosenbaum, P.R., and D. B. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, vol. 70, 41-55.
- Rubin, D.B. (1977), "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational Statistics*, vol. 2, 1-26.

DIFFERENCES-IN-DIFFERENCES AND SYNTHETIC CONTROLS

MM Chapter 5, MHE Chapter 5

- Abadie, A. (2005), "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies*, vol. 72, 1-19.
- Abadie, A. (2019), "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects," *Journal of Economic Literature* (forthcoming).
- Abadie, A., A. Diamond and J. Hainmueller (2010), "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, vol. 105, 493-505.
- Abadie, A. and J. Gardeazabal (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, vol. 93(1), 113-132.
- Athey, S., M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2018), "Matrix Completion Methods for Causal Panel Data Models. Working Paper 25132, National Bureau of Economic Research.
- Card, D. (1990), "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, vol. 44, 245-257.
- Card, D. and A.B. Krueger (1994), "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, vol. 84, 772-793.

- Duflo, E. (2001), "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, vol. 91, 795-813.
- Gruber, J. (1994), "The Incidence of Mandated Maternity Benefits," *American Economic Review*, vol. 84(3), 622-641.

REGRESSION DISCONTINUITY DESIGNS

MM Chapter 4, MHE Chapter 6

- Card, D., Lee, D., Pei, Z., and Weber, A. (2015). "Inference on causal effects in a generalized regression kink design." *Econometrica* 83(6).
- Calonico, S., Cattaneo, M., and Titiunik, R. (2014). "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6).
- Clark, D., and Martorell, P. (2014). "The signaling value of a high school diploma." *Journal of Political Economy* 122(2).
- Imbens, G., and Lemieux, T. (2008). "Regression discontinuity designs: a guide to practice." *Journal of Econometrics* 142(2).
- Imbens, G., and Kalyanaraman, K. (2012). "Optimal bandwidth choice for the regression discontinuity estimator." *Review of Economic Studies* 79(3).
- Lee, D. (2008). "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics* 142.
- Lee, D., and Lemieux, T. (2010). "Regression discontinuity designs in economics." *Journal of Economic Literature* 48(2).
- McCrary, J. (2008). "Manipulation of the running variable in the regression discontinuity design: a density test." *Journal of Econometrics* 142(2).
- Thistlethwaite, D., and Campbell, D. (1960). "Regression-discontinuity analysis: an alternative to the ex post facto experiment." *Journal of Educational Psychology* 51(6).
- Urquiola, M., and Verhoogen, E. (2009). "Class-size caps, sorting, and the regression discontinuity design." *American Economic Review* 99(1).

SELECTION MODELS AND POLICY EVALUATION

- Brinch, C., Mogstad, M., and Wiswall, M. (2017). "Beyond LATE with a discrete instrument." *Journal of Political Economy* 125(4).
- Carneiro, P., Heckman, J., and Vytlacil, E. (2010). "Evaluating marginal policy changes and the average effect of treatment for individuals at the margin." *Econometrica* 78(1).
- Heckman, J. (1979). "Sample selection bias as a specification error." *Econometrica* 47(1).
- Heckman, J., Urzua, S., and Vytlacil, E. (2006). "Understanding instrumental variables in models with essential heterogeneity." *Review of Economics and Statistics* 88(3).
- Heckman, J., and Vytlacil, E. (1999). "Local instrumental variables and latent variable models for identifying and bounding treatment effects." *Proceedings of the National Academy of Sciences* 96(8).

- Heckman, J., and Vytlacil, E. (2005). "Structural equations, treatment effects, and econometric policy evaluation." *Econometrica* 73(3).
- Kline, P., and Walters, C. (2016). "Evaluating public programs with close substitutes: the case of Head Start." *Quarterly Journal of Economics* 131(4).
- Kline, P., and Walters, C. (2019). "On heckits, LATE, and numerical equivalence." *Econometrica* 87(2).
- Olsen, R. (1980). "A least squares correction for selectivity bias." *Econometrica* 48(7).
- Vytlacil, E. (2002). "Independence, monotonicity and latent index models: an equivalence result." *Econometrica* 70(1).

BAYESIAN AND MACHINE LEARNING METHODS

- Abadie, A., and Kasy, M. (forthcoming). "The risk of machine learning." *Review of Economics and Statistics*.
- Angrist, J., Hull, P., Pathak, P., and Walters, C. (2017). "Leveraging lotteries for school value-added: testing and estimation." *Quarterly Journal of Economics* 132(2).
- Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). "Sparse models and methods for optimal instruments with an application to eminent domain." *Econometrica* 80(6).
- Belloni, A., Chernozhukov, V., and Wang, L. (2011). "Square-root LASSO: pivotal recovery of sparse signals via conic programming." *Biometrika* 98(4).
- Chetty, R., Friedman, J., and Rockoff, J. (2014). "Measuring the impacts of teachers II: teacher value-added and outcomes in adulthood." *American Economic Review* 104(9).
- Chetty, R., and Hendren, N. (2018). "Impacts of neighborhoods on intergenerational mobility II: county-level estimates." *Quarterly Journal of Economics* 133(3).
- Gilchrist, D., and Sands, E. (2016). "Something to talk about: social spillovers in movie consumption." *Journal of Political Economy* 124(5).
- Morris, C. (1983). "Parametric empirical Bayes inference: theory and applications." *Journal of the American Statistical Association* 78(381).
- Robbins, H. (1964). "An empirical Bayes approach to statistical decision problems." *Annals of Mathematical Statistics* 35(1).
- Varian, H. (2014). "Big data: new tricks for econometrics." *Journal of Economic Perspectives* 28(2).