

ECONOMETRIC DATA SCIENCE

MIT 14.32/14.320
Spring 2021

J. Angrist (angrist@mit.edu)
Juliette Fournier (jfournie@mit.edu)
Lauren Rice (lrice22@mit.edu)
Kiara Wahnschafft (kiaraw@mit.edu)
Luke Stewart (lrs22@mit.edu)

Econometric Data Science develops the knowledge and skills needed to understand empirical economic research and to plan and execute empirical projects. Topics include randomized trials, regression, instrumental variables, differences-in-differences, regression-discontinuity designs, and simultaneous equations models.

Prerequisites

Econometrics builds on the basic tools of probability and statistics, as taught in MIT 14.30 or similar. Our econometrics journey begins with a brief prob-stats refresher (just in case).

Course requirements

Eighty percent of success is showing up – Woody Allen

Classroom work:

Two lectures (TTH 10:30-12:00; E51-395 or remote) and weekly recitations (E51-395 or remote). Graduate and undergraduate recitations meet separately. We'll also schedule weekly computer labs throughout the semester.

As an incentive to show up, we take roll, marking attendance for those online as well as in-class. Classwork includes four pop quizzes and randomly assigned opportunities to share your understanding of econometric research with the class.

Written work:

1. Six graded problem sets and ungraded review problem sets at the beginning and end of the course. Problem sets include theoretical and data-analysis questions. The latter are to be answered using *Stata*, the lingua franca of applied econometrics. Classes focus on concepts and econometric applications rather than programming. Help for new *Stata* users is provided in computer lab and office hours.
2. A mandatory in-class midterm scheduled for class time TUESDAY APRIL 6, 2021.
3. A registrar-scheduled final during exam week.

14.320 students are asked to complete an ungraded capstone project. This is due Friday, May 28.

Grades

Grades are based on a point score, computed as follows: 30 points for problem sets (5 each), 20 points for the midterm, 30 points for the final, plus 25 additional points awarded as follows: up to 6 points for showing up (.25 for each class attended after the first; on-time arrival required), 4 each for 4 pop quizzes (absent or late counts as zero), and 3 for reading assignments scored via *Perusall*.

The P-set Deal

- The 6 graded problem sets are mandatory; solutions must be submitted on time (with *Stata* logs) to receive credit. Consult with classmates or your instructors if you get stuck, but solutions must be your own work.
- It's impossible to earn a passing grade on an exam-only basis: students with an average combined problem set grade below 50% through Problem Set 5 [CHANGE TO PSET 4 AS PSET 5 NOT GRADED BY DROP DATE] are ineligible to take the final and will be asked to drop the course.

Comportment

Like most things worth doing, econometrics requires focus and commitment. In this spirit, I ask you not to bring food to class (whether in-person or online) and to leave electronics shut off and put away except as needed for online participation.

Although lecture recordings will be posted, the class is offered synchronously. Online participants (this is everyone in the first two weeks of term) are asked to join us from a quiet, non-distracting location in full classroom dress. **Online participants should have Zoom video on and be prepared to participate as if attending in person.** Unless otherwise authorized, online attendance with video off counts as absent. Feel free to use zoom chat to fire off course-related questions to the group.

Texts and readings

We rely heavily on instructor notes, distributed in class. Our text is:

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

Occasional more advanced material comes from:

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

Journal articles and selected additional readings are posted on Canvas. We'll take advantage of Canvas's integration with Perusall, Piazza, and Gradescope, using Gradescope for exams as well as problem sets.

Our texts are inexpensive paperbacks, and may be rented for around \$10 per term. Many of our reading assignments are journal articles, posted on Canvas and available for free. Students will be asked to post (graded) comments and questions on these weekly via Perusall. We'll also randomly assign reading summaries in a form known as a peremptory pop quiz (PPQ).

Computer work

14.32 students can access cloud-based *Stata*. Please check with our TAs for info on set-up.

Teaching Assistants

Juliette Fournier is our Ph.D. student TA, responsible for course management and 14.320 recitations.

Lauren Rice is responsible for 14.32 recitations.

Kiara Wahnschafft is responsible for all things *Stata*.

Luke Stewart will offer tutorials for students interested in further review.

COURSE OUTLINE

What's it all for? Look ahead by reading:

MM, Introduction
MHE, Chapter 1

M.S. Kofoed, L. Gebhart, D. Gilmore, and R. Moschito, "Zooming to Class? Experimental Evidence on College Students' Online Learning During COVID-19," IZA DP No. 14356, May 2021.

S. Carter, K. Greenberg, and M. Walker, "The Impact of Computer Usage on Academic Performance: Evidence from a Randomized Trial at the United States Military Academy." *Economics of Education Review*, February 2017.

G. Burtless, "Are Targeted Wage Subsidies Harmful? Evidence from a Wage Voucher Experiment," *Industrial and Labor Relations Review* 39 (October 1989).

A. STATISTICAL TOOLS

Lecture Note 1: Expectation and Moments
Lecture Note 2: Sampling Distributions and Inference
Lecture Note 3: Confidence Intervals

MM, Chapter 1 Appendix

J. Angrist, D. Lang, and P. Oreopoulos, "Incentives and Services for College Achievement: Evidence from a Randomized Trial," *American Economic Journal: Applied Economics*, Jan. 2009.

J. Angrist, P. Oreopoulos, and T. Williams, "When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards," *J Human Resources* 49 (Summer 2014).

S. Woodbury, Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trials in Illinois, *The AER* (February 1987).

B. ANALYSIS AND INTERPRETATION OF RANDOMIZED TRIALS

Lecture Note 4: Causality and Potential Outcomes

MM, Chapter 1
MHE, Chapter 2

A. Aron-Dine, L. Einav, and A. Finkelstein, "The RAND Health Insurance Experiment Three Decades Later," *J. of Economic Perspectives* 27 (Winter 2013), 197-222.

R.H. Brook, et al., "Does Free Care Improve Adults' Health?," *New England J. of Medicine* 309 (Dec. 8, 1983), 1426-1434.

S. Taubman, *et al.*, "Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment," *Science*, Jan 2, 2014.

C. REGRESSION BASICS

Lecture Note 5: Intro to Multivariate Regression

Lecture Note 6: Understanding Multivariate Regression – the OVB Formula

Lecture Note 7: Regression Inference

-- approximate midterm date --

Lecture Note 8: Residuals, Fitted Values, and Goodness of Fit

MM, Chapter 2

MHE, Sections 3.1 (through 3.1.3), 3.2 (through 3.2.2), and 3.4.3

ALO (2009) and AOW (2014) listed under Part A.

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, 1491-1529.

S.B. Dale and A.B. Krueger, “Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data,” *The Journal of Human Resources*, vol. 49, no. 2, Spring 2014).

J. Mountjoy and B.R. Hickman, “The Returns to College(s): Estimating Value-Added and Match Effects in Higher Education,” Becker-Friedman Working Paper No. 2020-08, February 2020.

D. USING MULTIVARIATE REGRESSION

Lecture Note 9: Modeling with Multivariate Regression Models

MM, Chapter 2 Appendix

MHE, Section 3.1.4

A. Krueger, “How Computers Have Changed the Wage Structure: Evidence from Micro Data,” *Quarterly Journal of Economics* 108[1], February 1993, 33-60.

J. DiNardo and J.S. Pischke, “The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?,” *The Quarterly Journal of Economics* 112[1], February 1997, 291-303.

D. Autor, L. Katz, and A. Krueger, “Computing Inequality: Have Computers Changed the Labor Market?,” *The Quarterly Journal of Economics* 113 (November 1998).

Lecture Note 10: Standard Standard Error Issues

MM, Chapter 2 Appendix

MHE, Section 3.4.1

K. Graddy, “Testing for Imperfect Competition at the Fulton Fish Market,” *The RAND Journal of Economics* 26, Spring 1995, 75-92.

K. Graddy, "The Fulton Fish Market," *Journal of Economic Perspectives* 20, Spring 2006, 207-20.

A. Krueger, "Experimental Estimates of Education Production Functions," *Quarterly Journal of Economics* 115(2), May 1999, 497-532.

J. Angrist and V. Lavy, "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial," *The AER* 99, September 2009.

E. OMITTED VARIABLES SOLUTIONS

Lecture Note 11: Instrumental Variables and 2SLS for Omitted-Variable Problems

MM, Chapters 3 and 6
MHE, Sections 4.1 and 4.6.1

J. Angrist, "Lifetime Earnings and the Vietnam-Era Draft Lottery: Evidence from Social Security Administrative Records," *American Economic Review* 80, June 1990.

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

J. Angrist and W. Evans, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review* 88(3), June 1998.

J. Gelbach (2002), "Public Schooling for Young Children and Maternal Labor Supply," *American Economic Review* 92 (March 2002).

J. Angrist, *et al.*, "Who Benefits from KIPP?," *J. of Policy Analysis and Management*, Fall 2012.

J. Angrist, S. Cohodes, S. Dynarski, P. Pathak, and C. Walters, "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics* 34 (April 2016).

J. Angrist, V. Lavy, and A. Schlosser, "Multiple Experiments for the Quantity and Quality of Children," *Journal of Labor Economics* 28, October 2010.

A. Finkelstein, *et al.*, "The Oregon Health Insurance Experiment: Evidence from the First Year," *The QJE* 127 (August 2012).

W. Evans and R. Schwab, "Finishing High School and Starting College: Do Catholic Schools Make a Difference," *The QJE* 110 (November 1995).

A. Abdulkadiroglu, P. Pathak, and C. Walters, "Free to Choose: Can School Choice Reduce Achievement?," *AEJ: Applied Economics* 10 (2018).

Lecture Note 12: Panel Data, Fixed Effects, and IV for Measurement Error

MM, Chapter 6.

O. Ashenfelter and A. Krueger, “Estimates of the Economic Returns to Schooling from a New Sample of Twins,” *The AER* 84 (December 1994).

O. Ashenfelter and C. Rouse, “Income, Schooling, and Ability: Evidence from a New Sample of Twins,” *The QJE* 113 (February 1998).

Lecture Note 13: Doing Diffs-in-diffs

MM, Chapter 5
MHE, Section 5.2

G. Richardson and W. Troost, “Monetary Intervention Mitigated Banking Panics During the Great Depression: Quasi-Experimental Evidence from a Federal reserve District Border, 1929-1933,” *The JPE* 117 (2009).

D. Card, “The Impact of the Mariel Boatlift on the Miami Labor Market,” *Industrial and Labor Relations Review* 43 (January 1990), 245-257.

D. Card and A. Krueger, “Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania,” *American Economic Review* 90 (1994), 1397-420.

C. Carpenter and C. Dobkin, “The Minimum Legal Drinking Age and Public Health,” *The Journal of Economic Perspectives* 25 (2011), 133-156.

T. Deryugina and D. Molitor, “Does Where You Die Depend on Where You Live? Evidence from Hurricane Katrina,” *The AER* 110 (November 2020).

F: MORE 'METRICS MAGIC (TIME PERMITTING)

Lecture Note 14: RD in Action

MM, Chapter 4
MHE, Chapter 6

D. Lee, “Randomized Experiments from Non-random Selection in U.S. House Elections,” *J. Econometrics* 142 (2008), 675-97.

C. Carpenter and C. Dobkin, “The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the MLDA,” *American Economic Journal: Applied Economics* 1 (2009), 164-182.

A. Abdulkadiroglu, *et al.*, “The Elite Illusion: Achievement Effects at Boston and New York Exam Schools,” *Econometrica*, 2014.

J. Angrist and V. Lavy, “Using Maimonides Rule to Estimate the Effect of Class Size on Earnings,” *The QJE*, May 1999.

J. Angrist, V. Lavy, J. Leder-Luis, A. Shani, “Maimonides Rule Redux,” *AER: Insights* 1 December 2019.

D. Almond, J. Doyle, A. Kowalski, and H. Williams, "Estimating the Marginal Returns to Medical Care: Evidence from At-Risk Newborns," *The QJE* 125(2), 2010, 591-634.

A. Barreca, M. Guildi, J. Lindo, and G. Waddell, "Saving Babies? Revisiting the Effect of Very Low Birthweight Classification," *The QJE* 126(4), November 2011, 2117-2123.

S. Cohodes and J. Goodman, "Merit Adi, College Quality, and College Completion: Evidence from Massachusetts' Adams Scholarship," *American Economic Journal: Applied Economics* 6 (2014), 251-285.

Lecture Note 15: Simultaneous Equations Models

J. Angrist, G. Imbens, K. Graddy, "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish," *Review of Economic Studies* 67[3], July 2000, 499-257(29).