Econometrics

MIT (14.32)
Spring 2015

J. Angrist (angrist@mit.edu)
Luu Nguyen (hqn@mit.edu)
Stephanie Cheng (sdcheng@mit.edu)

Our aim is to help you to understand modern applied econometric methods and to foster the skills needed to plan and execute your own empirical projects. Topics include randomized trials, regression, differences-in-differences, instrumental variables, regression-discontinuity designs, and simultaneous equations models. We study many examples and do a fair amount of number crunching ourselves.

Prerequisites

Students should be familiar with basic concepts in probability and statistics. The course begins with a 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; E51-057) and a weekly recitation (F 9:00 E51-149).

As an incentive to show up, we take roll. There are also four (4) in-class pop quizzes.

Other work:

You'll finish the course equipped with a workman's familiarity with the tools of probability and statistics, facility with data handling and statistical programming, and—hopefully—a good understanding of the models and methods of applied econometrics. That's a lot of ground to cover, so plan your time accordingly. There are 5 graded problem sets and ungraded review problem sets at the beginning and end of the course. The problem sets have both analytical and computer-exercise components. *Stata* is our default programming language for problem sets and in recitation. Classes focus on concepts and econometric applications. Help for new *Stata* users will be given in recitation and by our grader. We'll have an in-class midterm and a final during exam week.

Our (mandatory) in-class midterm is scheduled for March 19, 2015. There is no make-up or conflict midterm.

Grades

Showing up is 80% of success, but it's only 8% of your grade. Grades are computed as follows: a total of 130 points, 30 points for problem sets (6 points each), 30 points for the midterm, 40 points for the final, and 30 bonus points awarded as follows:

Up to 10 points for attendance (.5 for each class attended up to 20; on-time arrival required)
5 each for 4 pop quizzes (absent or late counts as zero).

There is a 6th ungraded review problem set.

Graded problem sets are mandatory and solutions must be submitted on time to receive credit. *Stata* logs are to be submitted with solution sets. A grade of 75% or better on at least 4 problem sets is required in order to be eligible to take the final. **Consult with classmates on problem sets if you get stuck, but solutions must be your own work.**

Comportment

This course is ill-suited to strategically adept effort-minimizing time-managers in the grand MIT tradition. It's more like love: success follows clear expressions of commitment and desire. Econometrics, like most things worth doing, requires focus and attention.

In this spirit, I ask you not to bring food to class and to leave electronics and other toys shut off and put away (this prohibition includes but is not limited to: laptops and tablets; ipods; phones; Wii, Xbox, or Playstation consoles; inflatable love dolls). Airplane mode not allowed.

Texts and readings

We rely heavily on:

Instructor notes, distributed in class.

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

Problems and additional readings come from:

J. Stock and M. Watson, *Introduction to Econometrics* (3rd ed.), Addison Wesley, 2011 (SW).

For those who want to dig deeper:

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

Journal articles and selected additional readings are posted on our Stellar web site (http://stellar.mit.edu/S/course/14/sp15/14.32/).

Computer work

For the purposes of this course, you'll have access to cloud-based *Stata* to run on your own laptop or the computer of your choice. Please check with our TAs for info on set-up.

Course outline for 14.32

The Big Picture

We start with a stats review based on my notes. Look ahead by reading:

MHE, Chapter 1
MM, Intro

A. Statistical Tools

Lecture Note 1: Expectation and Moments

MM, Chapter 1 Appendix
SW, Chapter 2

B. Review of Statistical Inference

Lecture Note 2: Sampling Distributions and Inference
Lecture Note 3: Confidence Intervals

MM, Chapter 1 Appendix
SW, Chapter 3

C. Analysis and Interpretation of Randomized Trials

Lecture Note 4: Causality, Experiments, and Potential Outcomes

MM, Chapter 1

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.

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.

Lecture Note 5: Life in Asymptopia

D. Regression Basics: Why and How?

Lecture Note 6: Bivariate Regression and the CEF
Lecture Note 7: Introduction to Multivariate Regression
Lecture Note 8: Multivariate Regression (cont.) – Omitted Variables, Short vs. Long

Lecture Note 9: Sampling Distribution of Regression Estimates
Lecture Note 10: Residuals, Fitted Values, and Goodness of Fit

MM, Chapter 2
SW, Chapters 4-7 and 17.1-17.4
MHE, Sections 3.1 (through 3.1.3), 3.2 (through 3.2.2), and 3.4.3

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," NBER Working Paper 17159 (June 2011; forthcoming in *The Journal of Human Resources*, 2014

-- approximate midterm date --

E. Using Multivariate Regression

Lecture Note 11: Dummy Variables, Interactions, F-Tests

MM, Chapter 2 Appendix
SW, Chapters 8-9
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.

F. Inference Problems in Asymptopia; Heteroskedasticity and Serial Correlation

Lecture Note 12: Asymptotic Distribution of Regression Estimates
Lecture Note 13: Heteroskedasticity, GLS, and Linear Probability Models (Worrying about SEs, Part I)
Lecture Note 14: Serial Correlation and Clustering (Worrying About SEs, Part II)

MM, Chapter 2 Appendix
SW, Chapters 14.1-14.3, 15.4, 17.5
MHE, Section 3.4.1

G. Instrumental Variables

Lecture Note 15: Instrumental Variables and Two-Stage Least Squares for Omitted-Variables Problems

MM, Chapters 3 and 6
SW, Chapter 12, 13.5-13.7, and Appendices to Chapter 13
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*, June 1990.

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

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

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

H. Regression Discontinuity Designs

Lecture Note 16: RD in Action

MM, Chapter 4
SW, Section 13.4-13.5
MHE, Chapter 6

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.

I: Differences-in-Differences and State Policy Experiments

Lecture Note 17: DD and the Min

MM, Chapter 5
SW Chapters 10 and 13.1-13.4
MHE, Section 5.2

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.

D. Card, "Using Regional Variation to Measure the Effect of the Federal Wage," Industrial and Labor Relations Review (1992) 46, 22-37.

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

J. Simultaneous Equations Models

Lecture Note 18: 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).

MIT 14.32 **Spring 2015**
J. Angrist

Course Outline and Study Guide

Part I

A. Statistical tools

1. Expectation and moments

B. Review of statistical inference

2. Sampling distributions and inference
3. Confidence intervals

C. Analysis and interpretation of randomized trials

4. Causality, experiments, and potential outcomes; statistical and causal inference
5. Life in asymptopia (statistical inference in practice)

D. Regression basics

6. Bivariate regression and the conditional expectation function
7. Multivariate regression, causality, and control; anatomy of multivariate regression coefficients
8. Omitted variables, short vs. long regressions
9. Sampling distribution of regression estimates (origin of regression SEs)
10. Residuals, fitted values, and goodness of fit

Part II

E. Using multivariate regression

11. Dummy variables and interactions; testing linear restrictions using F-tests

F. Inference Problems

12. Regression in Asymptopia
13. Heteroskedasticity—Who cares? Weighted least squares; the linear probability model, GLS vs. fix-the-standard-errors approaches to het.
14. Serial correlation in time series—Who cares? Serial correlation fix-ups and tests

G. Instrumental variables (IV)

15. Using IV and 2SLS to solve omitted-variables problems

H. Regression Discontinuity Designs

16. RD in Action

I. Differences-in-Differences

17. DD, Regression DD, minimum wage effects

J. Simultaneous Equations Models

18. Simultaneous equations models: motivation and identification

The use of structural models
Simultaneous equations bias and the identification problem
Structural equations and reduced form equations
Indirect least squares (ILS)
IV for the SEM
2SLS and SEM identification