Econometric Data Science develops the theoretical knowledge and applied 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) and a weekly recitation (14.32: F 11-12 room TBD; 14.320 Th 4-530, room TBD). The 14.320 recitation includes a research workshop described below. We’ll also offer optional computer labs and tutorials.

As an incentive to show up, we take roll and credit attendance. Classwork includes two pop quizzes and randomly assigned opportunities to share your understanding of econometric research with the class.

Problem sets and exams:

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. Two pop quizzes and a mandatory in-class midterm on Thursday, March 17, 2022.
3. A registrar-scheduled final during exam week.

14.320 students may also submit an ungraded capstone project for instructor feedback. This is due Friday, May 20, 2022.

Grades are based on a point score, computed as follows:

- 14.32 - 30 points for problem sets (5 each), 20 points for the midterm, 32 points for the final, plus 18 additional points awarded as follows: up to 5 points for showing up (.2 for each class attended after the first; on-time arrival required), 4 each for 2 pop quizzes (absent or late counts as zero), and 5 for reading assignments scored via Perusall.
• 14.320 - 24 points for problem sets (4 each), 20 points for the midterm, 32 points for the final, plus 24 additional points awarded as follows: up to 5 points for showing up (.2 for each class attended after the first; on-time arrival required), 4 each for 2 pop quizzes (absent or late counts as zero), and 3 for reading assignments scored via Perusall.
  o As part of weekly recitation, 14.320 students participate in a research workshop where they give and discuss their classmates’ 20-minute presentations of articles from our reading list. Workshop participants earn 6 points for presentation and 2 for participation.

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.
• 14.32/320 can’t be passed on an exam-only basis: students with an average problem set grade under 50% through Pset 4 are ineligible to take the final and will be asked to drop the course.

Comportment

• Econometric mastery requires focus and commitment. In this spirit, I ask you not to bring food to class, to leave electronics shut off and put away, and be prepared to participate in class discussion.
• Thursday classes begin with two quick questions on our readings (2QQs), randomly assigned.

Texts and readings


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 articles weekly via Perusall.
• Most classes begin with two quick questions (2QQs) about readings, with respondents chosen by random assignment.

Computer work

14.32/320 students use our departmental *Stata* license. Please check with our TAs for info on set-up.

Teaching Assistants

Andrea Manera is our head TA, responsible for course management and Friday recitations.
Viola Corradini is TA in charge of 14.320 and the research workshop.
Isabel Munoz is responsible for all things *Stata*.
Grace Chuan 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


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


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


C. Regression Basics

Lecture Note 5: Intro to Multivariate Regression
Lecture Note 6: Understanding Multivariate Regression – Causality and Control, 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.


D. Using Multivariate Regression

Lecture Note 9: Modeling with Multivariate Regression Models

MM, Chapter 2 Appendix
MHE, Section 3.1.4


E. Omitted Variables Solutions

Lecture Note 11: Instrumental Variables and 2SLS

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


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

MM, Chapter 6.


Lecture Note 13: Doing Diffs-in-diffs

MM, Chapter 5
MHE, Section 5.2


F: More ‘Metrics Magic (time permitting)

Lecture Note 14: RD in Action

MM, Chapter 4
MHE, Chapter 6


Lecture Note 15: Simultaneous Equations Models