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:


Occasional more advanced material comes from:


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 Wahnschaft 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


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 – 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


Lecture Note 10: Standard Standard Error Issues

MM, Chapter 2 Appendix
MHE, Section 3.4.1


**6. 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


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