The University of Texas at Austin
ECO 394M (Master's Econometrics)
Prof. Jason Abrevaya
Fall 2023

SYLLABUS

ECO 394M is an introductory Master's level course in econometrics. The course will cover a lot of ground, focusing on both the theoretical underpinnings of the econometric methods and the application of these methods to actual data. Topics will include (as time permits): linear regression model, asymptotic (large-sample) and bootstrap inference, model specification, hypothesis testing, instrumental variables (IV), general method of moments (GMM) estimation, quantile regression, panel data, nonlinear models (binary choice, count data, etc) and MLE estimation, time series models.

Contact information: Please contact me by email (abrevaya@austin.utexas.edu) with any questions about the course. I will hold weekly office hours (time/day on Canvas) and probably additional office hours prior to exams.

Teaching assistant: Nathan ("Nate") Hattersley (nhattersley@utexas.edu) will lead a weekly review section and hold weekly office hours.

Prerequisities: Knowledge of probability and statistics, at the level of ECO 394D, and multivariable calculus and linear algebra, at the level of ECO 394C, will be assumed.

Textbook: Although the lecture notes will contain the crucial course material, students will be asked to read *Introductory Econometrics: A Modern Approach* by Jeffrey Wooldridge. This textbook is intended for an undergraduate econometrics course, and we will cover the material in class at a more advanced and rigorous level. The textbook provides excellent intuition behind most of the topics that we will cover in class and, therefore, serves as very nice complement to the lectures. The textbook is now in its 7th Edition, but students may also use the earlier (cheaper) 5th or 6th Editions. Please note that we will also cover topics in lecture (for example, GMM and the bootstrap) that are not covered in the textbook. For interested students, more advanced treatments of some of the course's topics can be found in *Econometric Analysis of Cross Section and Panel Data* by Jeffrey Wooldridge, *Econometrics* by Bruce Hansen, and *Time Series Analysis* by James Hamilton (for time-series topics).

Course materials: All course materials and non-textbook readings will be posted on the Canvas site. Lecture notes will be posted as pdf files prior to class. Annotated notes, with handwriting from lecture, will be posted a few days after the completion of each part of the notes.

Grading: Grades will be based on a midterm (20%), problem sets (45%), and a final (35%). The final exam will be held on **Monday, December 11th at 8:00-11:00am**. The midterm date/time will be announced when it is known.

Software: Students are encouraged to use Stata for the empirical homework exercises and also to "practice" applying the econometric methods. An excellent set of Stata resources is maintained by UCLA at https://stats.idre.ucla.edu/stata/ (including their "web books" on regression and logistic regression at https://stats.idre.ucla.edu/stata/webbooks/). The Stata website also has a list of useful links at https://www.stata.com/links/resources-for-learning-stata/. If you prefer, you may use another programming language (*R* or *Python*) for the problem sets, though it will probably require more time on your part and you'll still need to be somewhat familiar with Stata output for the exams.

Problem sets: You are encouraged to work with other students, but each student must write up their own solutions. Please include the relevant computer output (i.e., commands used to do the analysis and the output itself). Do not use AI tools (e.g. ChatGPT) or solutions from elsewhere to assist with the problem sets, as you will not learn the material and, in the latter case, will be considered plagiarism.

List of topics:

(W indicates suggested background reading in Wooldridge; * indicates a topic that is largely not covered in the book.)

- Linear regression: model and estimation (W 1, 2, 3, 6.1-6.2, 7.1-7.4)
  - Model: structural versus statistical views
  - Model parameters and interpretation
  - Model specification: polynomials, interactions, indicator variables, functional form (logs)
  - OLS estimation: fitted values, residuals, R-squared
- Statistical inference (W 4, 5, 6.4)
  - Asymptotic (large-sample) theory: consistency, asymptotic normality, variance formulas, standard errors, confidence intervals, functions of parameters, multicollinearity, forecasts
  - Hypothesis testing: z-test (asymptotic "t" test), Wald test (asymptotic "F" test)
  - Bootstrap inference
- Other issues (W 8.1-8.4, 9.4, 16.1-16.2)
  - Heteroskedasticity: testing, weighted least squares (sidebar on nonlinear least squares), forecasts
  - Serial correlation: more later
  - Failures of exogeneity

- *Quantile regression (W 9.6)
  - Conditional quantile models
  - Least absolute deviations (LAD) and general quantile estimation
- *Generalized method of moments (GMM) estimation (W 15)
  - Simple examples
  - Instrumental variables (IV) model and estimation
- Panel data (W 13, 14)
  - Fixed effects model: strict exogeneity, within and first-difference estimation
  - Failure of strict exogeneity: dynamic models, *feedback effects, *GMM solutions
- Nonlinear models (W 7.5, 17)
  - Binary-choice model: MLE, partial effects
  - Count data model
  - Other models: censored data, corner-solution data
- Time series (as time permits) (W 10, 11, 12, 18)
  - Simple model features: finite distributed lags, trends, seasonality
  - Autoregressive models
  - Other issues: Stationarity vs non-stationarity, cointegration