PPD 716: Econometrics for PPM I

Nicolas Duquette

Fall 2020

Monday – 2PM to 5:20 PM
Location: Online-only

Instructor: Nicolas Duquette
Office Hours: Online, 11am-noon on Friday
Email: nduquett@usc.edu

Course Description

This course is designed to expand students' knowledge of econometric methods. In particular, the course will develop econometric reasoning and skills through practical applications. The course builds on the foundations of PPD558 and PPDE668 and prepares students for Econometrics for Policy, Planning and Management II. The course is appropriate for Ph.D. students in Price School and other students in the social sciences with the appropriate econometrics background.

Learning Objectives

Upon completion of this course, students should be able to:

- implement a wide range of microeconometric techniques
- critically assess applied quantitative studies
- identify empirical strategies appropriate for a given research question and data structure
- perform a variety of data preparation, programming and analytic tasks

Pandemic Adaptation

This course has been designated as online-only for Fall 2020 because of the Sars-CoV-2 pandemic. Opportunities to engage with the instructor and other students fall into four types:

1. *Lectures* – Lectures will be posted as recorded videos to the Google Drive, along with PDFs of lecture slides. Students are expected to watch all the lectures before each week's class meeting. Lectures are divided into multiple shorter videos by topic for, hopefully, less painful viewing.
2. *In-class exercises and discussion* – we will meet synchronously via Zoom at the scheduled time (2pm Pacific time on Mondays) to work through in-class exercises and have open discussion of any questions you have from the lecture material. Because the lecture content will have already been viewed asynchronously, these Zoom meetings will likely be substantially shorter than the full three hours and twenty minutes allocated.
3. *Office hours* – We will have drop-in "office hours" by Zoom Fridays 11am to noon, US Pacific Time. These won't literally involve my office. Instead I will be on Zoom and you are welcome to log in with any questions or topics you'd like to chat about outside of our Monday afternoon meetings. I am also happy to correspond about the class by email or to set up one-on-one video calls by appointment.
4. *Assignments* – All grading in this class is based on assignments described below. You are strongly encouraged to work in groups on these. I will facilitate groupwork for the first assignment by assigning partners.

This syllabus is subject to change as the public health environment changes. Please be patient and understanding with any on-the-fly adjustments.

Similarly, I promise to be accommodating and understanding if the pandemic impinges on your ability to meet class obligations. Just talk to me.

Description and Assessment of Assignments

There will be seven problem sets in this class, each requiring students to examine and replicate a portion of a published paper or otherwise apply class techniques to real data. You are allowed and encouraged to work in teams of up to four people on these problem sets. Answers to each problem set must be submitted in PDF format with the inclusion of supplementary files in a compressed directory format (.TGZ, .7Z or .ZIP). Each team needs to submit only one copy of its problem set by email to `nduquett@usc.edu`. Each problem set is worth 10 points.

Problem Set Due Dates

Problem Set	Topic	Due Date
1	Warmup	August 31
2	Difference-in-difference	September 14
3	Regression discontinuity	September 28
4	Synthetic control	October 12
5	Event study	October 26
6	Multilevel models	November 2
7	Student's choice	November 24

I will also assign two referee reports. You will read a real research paper and critique its econometric methodology. Each should take about 200-400 words. You may work in teams of two. Cumulatively, these referee reports are worth thirty points. Further instructions and resources are available in the Reports folder of the Google Drive.

Your reports are due on the following dates. Submit one copy per group in PDF format only to `nduquett@usc.edu`

- White and Wagner, Teacher Strikes – September 21
- TBD – November 24

Grading by points

Total Points	Letter Grade
$\geq$ 93 out of 100	A
90-92	A-
87-89	B+
83-86	B
80-82	B-
$\leq$79	$\leq$ C+

Technology

You are expected to attend each virtual each class meeting with a computer equipped with Stata. Stata is statistical software commonly used in policy research, and the most-used econometrics software within the Price School.

We realize that attending classes online and completing coursework remotely requires access to technology that not all students possess. If you need resources to successfully participate in your

classes, such as a laptop or internet hotspot, you may be eligible for the university's equipment rental program. To apply, please submit an application at `https://studentbasicneeds.usc.edu/resources/technology-assistance/`. The Student Basic Needs team will contact all applicants in early August and distribute equipment to eligible applicants prior to the start of the fall semester.

You can purchase Stata at StataCorp's web site.[1] Stata/IC (the cheapest but quite limited option) should be sufficient for almost everything we do in this course. You can access virtual machines with Stata/SE (the fancy version) installed at `https://cloudapps.usc.edu`. See also the list of software available to USC students at `https://software.usc.edu/`.

We will also touch on Python, a general-purpose programming language increasingly used for data analytics, particularly machine learning methods (`http://www.python.org`). Stata version 16 integrates Python directly into the do-file environment, making Python and Stata complementary skills.

If you are already very familiar with Stata, you are welcome to try using a different software system to complete the assignments, with the warning that the instructor does not provide support for other software. The onus is on the student to successfully complete the assignments. The statistics software R is open-source, free of charge, and widely used, and would be a good choice if you want to learn a second econometric toolkit. You can obtain R at `https://www.r-project.org/`. You can also use Python for assignments designed for Stata.

Course Web Sites

Lecture videos, lecture slides, assignments, data sets, and other useful resources will be posted to the class Google Drive. You must be logged into your USC account to access the materials.

Recorded Zoom meetings and grades will be posted to the class Blackboard site.

USC **prohibits sharing of any synchronous and asynchronous course content** outside of the learning environment. Sharing of course materials will lead to academic integrity sanctions. Specifically, SCampus Section 11.12(B) states

Distribution or use of notes or recordings based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study is a violation of the USC Student Conduct Code. This includes, but is not limited to, providing materials for distribution by services publishing class notes. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relationship to the class, whether obtained in class, via email, on the Internet or via any other media. (See Section C.1 Class Notes Policy).

[1] `https://www.stata.com/order/new/edu/gradplans/student-pricing/`

Synchronous session recording notice

Zoom class meetings will be recorded and provided to all students asynchronously via Blackboard. SCampus Section 11.12(C) prohibits student recording of class meetings. Do not redistribute class recordings.

Office hours will not be recorded.

Readings Overview

Each class meeting has an associated list of suggested readings. No specific reading is required, but all have potential to be useful.

Where possible, I have placed readings on digital reserve. I definitely recommend studying from reserve copies before buying to make sure you have the right text for your future research needs. Unfortunately, because of the pandemic, it is likely that reserves will not be physically open this semester and some resources will not be available virtually.

- Cunningham, S. (2021). *Causal Inference: The Mixtape*. Forthcoming from Yale University Press, New Haven. Version 1.8 is free and available at `http://scunning.com/mixtape.html`. Approachable overview of causal inference with workhorse estimators in economics. Copious Stata examples. Has an accompanying Spotify playlist.
- Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press. Works through the matrix algebra for all major econometric techniques, with examples. LOC HB172.C343.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. Accessible yet rigorous exploration of most common causal identification strategies. LOC HB139.A54.
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press. Deep dive into panel methods. LOC HB139.W663.
- Long, J. S. and Freese, J. (2014). *Regression Models for Categorical Dependent Variables Using Stata*. Stata Press, College Station, TX, third edition. Excellent coverage of binary, categorical, and count dependent variable models. Includes authors' own code for straightforward calculation of marginal effects. Does not cover new features in Stata v16, such as `cmset` commands. LOC QA 278.2 L848r

- Morgan, S. L. and Winship, C. (2015). *Counterfactuals and Causal Inference*. Cambridge University Press, second edition. Emphasis on application of causal inference to econometric methods. The authors are sociologists and the examples come from both sociology and economics. LOC H62.M646.
- Hansen, B. (2020). *Econometrics*. University of Wisconsin. Free and available at `https://www.ssc.wisc.edu/~bhansen/econometrics/`. An econometrics textbook used in the first year of the economics PhD at Wisconsin. Useful if you would like to see the probability-theoretic underpinnings of particular estimators.

Course Schedule: A Weekly Breakdown

Below is a list of topics by weeks and associated methodological readings. In addition to these readings, I will present applications of each technique from real research in class. Bibliographies for all examples will be included in the reference section of the slide decks.

August 17, 2020

Description, prediction, and explanation; Statistics review; Potential outcomes; Randomized social experiments; Coding workflows and research best practice

- Angrist and Pischke, chapter 2.
- Cameron and Trivedi, chapter 2.
- Morgan and Winship, chapters 1 and 2.
- Cunningham, "Potential outcomes causal model"
- Gentzkow, M. and Shapiro, J. M. (2014). *Code and Data for the Social Sciences: A Practitioner's Guide*. Available at `https://web.stanford.edu/~gentzkow/research/CodeAndData.pdf`
- `https://xkcd.com/2180/`

August 24, 2020

Least-squares regression; Probit and logistic regression

- Cameron and Trivedi, chapters 4 and 14.

- Cunningham, "Properties of regression"
- Angrist and Pischke, chapter 3.
- Hansen, chapter 25, "Binary choice."

August 31, 2020

Propensity score matching and weighting; 2x2 Difference-in-differences; Difference-in-difference-in-differences

- Cunningham, "Differences-in-differences" and "Matching and subclassification"

September 7, 2020

LABOR DAY (NO CLASS)

September 14, 2020

Causal graphs; Instrumental variables

- Cunningham, "Directed acyclical graphs" and "Instrumental Variables"
- Morgan and Winship, chapters 3 and 9.
- Pearl, J. (2009). *Causality*. Cambridge University Press, second edition. Pearl's *magnum opus*. USC has access to a cumbersome (but free-of-charge) e-book version through the university libraries.
- Pearl, J., Glymour, M., and Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. Wiley, first edition. Again, USC has access to a cumbersome (but free-of-charge) e-book version through the university libraries.
- Angrist and Pischke, chapter 4.

September 21, 2020

Regression discontinuity; Differences-in-regression discontinuity; Notches and kinks

- Cunningham, "Regression Discontinuity"
- Angrist and Pischke, chapter 6.

- Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2018). *A Practical Introduction to Regression Discontinuity Designs*. Forthcoming, Cambridge University Press. Includes Stata and R code throughout. Two volumes, available as preprint PDFs at `https://sites.google.com/site/rdpackages/home`.
- Skovron, C. and Titiunik, R. (2015). A practical guide to regression discontinuity designs in political science. *American Journal of Political Science*, 2015:1–36

September 28, 2020

Time series basics; Interrupted time series; Introduction to panel data; Synthetic control

- Pickup, M. (2014). *An Introduction to Time Series Analysis*. SAGE Publications
- Becketti, S. (2013). *Introduction to Time Series using Stata*. Stata Press, College Station, Texas, chapters 2 and 5.
- Morgan and Winship, chapter 11.
- Cunningham, "Synthetic control"

October 5, 2020

Fixed effects; Panel diff-in-diff; Event study models

- Cunningham, "Panel data."
- Cameron and Trivedi, chapters 21–22.
- Angrist and Pischke, chapter 5.
- Hatfield, L. and Zeldow, B. (2019). Difference-in-differences. Free guide online at `https://diff.healthpolicydatascience.org/`. An approachable exposition of DD methods with illustrative gifs.
- Goodman-Bacon, A. (2019). Difference-in-differences with variation in treatment timing. Working Paper 25018, National Bureau of Economic Research
- Goodman-Bacon, A., Goldring, T., and Nichols, A. (2019). `bacondecomp`: Stata module for decomposing difference-in-differences estimation with variation in treatment timing. Statistical Software Components S458676, Boston College Department of Economics

October 12, 2020

Standard errors; Clusters; Bootstrap; Randomization inference

- Cameron and Trivedi, chapters 11 and 24.
- Angrist and Pischke, chapter 8.
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427
- Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? Working Paper 24003, National Bureau of Economic Research
- Heß, S. (2017). Randomization inference with Stata: A guide and software. *Stata Journal*, 17(3):630–651
- Kerwin, J. (2017). Randomization inference vs. bootstrapping for p-values. Blog post, `https://jasonkerwin.com/nonparibus/2017/09/25/`
- Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249
- Correia, S. (2014). `reghdfe`: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. Statistical Software Components S457874, Boston College Department of Economics, revised 10 Mar 2019
- Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., and Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using `boottest`. *The Stata Journal*, 19(1):4–60
- Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372

October 19, 2020

Generalized least squares; Random effects; Multi-level modeling; Dynamic panel models; Linear systems; Seemingly unrelated regression; Simultaneous equation systems

- Rabe-Hesketh, S. and Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata*. Stata Press, College Station, TX, third edition. Needs an update to more recent Stata syntax, but extremely thorough exposition of the method and its applications.

- Wooldridge
- Angrist and Pischke, chapter 5.
- Cameron and Trivedi, chapters 2.4-2.5 and 6.
- Hansen, chapter 11.8.

October 26, 2020

Censored and truncated regression; Sample selection models; Hurdle models; Measurement error

- Cameron and Trivedi, chapters 26, 27.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161

November 2, 2020

Duration and survival analysis; Count models; Spatial regression

- Cameron and Trivedi, chapters 17, 20.
- Long and Freese, chapter 9.
- Cleves, M., Gould, W. W., and Marchenko, Y. V. (2016). *An Introduction to Survival Analysis Using Stata*. Stata Press, College Station, TX, revised third edition. LOC QA276.2.C54.
- Hilbe, J. M. (2014). *Modeling Count Data*. Cambridge University Press. Many examples of Stata and R code.
- Cameron and Trivedi, chapter 26.
- Ward, M. D. and Gleditsch, K. S. (2008). *Spatial Regression Models*. SAGE Publications. Some examples of code using R

November 9, 2020

Heterogeneity; quantile regression; distribution regression

- Cameron and Trivedi, chapter 4.6.

- Hao, L. and Naiman, D. Q. (2007). *Quantile Regression*. Sage Publications

Our modified semester ends November 13; the following class meetings will be posted to the Google Drive with an optional discussion section to be scheduled.

Lecture #13

Basic machine learning; Lasso regression

- StataCorp (2019). *Stata Lasso Reference Manual*. Stata Press.
- Burkov, A. (2019). *The Hundred-page Machine Learning Book*. http://themlbook.com/. Available as a pay-what-you-want e-book. Short but clear.
- Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, second edition. Look at this if you need to master a particular ML tool.
- Belloni, A., Chernozhukov, V., and Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50

Lecture #14

Ordered logit; Multinomial logit; Decision trees; Random forests

- Cameron and Trivedi, chapter 15.
- Long and Freese, chapters 5–8.
- Corp., S. (2019). *Stata Choice Models Reference Manual*. Stata Press, release 16 edition. `https://www.stata.com/manuals/cm.pdf`.
- Burkov.

Parental Obligations

As long as you take care not to disrupt other students, you may bring your child to class on school holidays or during child care disruptions.

If your child joins us, please sit near the door so you can step out with any young ones who get fussy, or please ensure your older child understands that they need to focus on a quiet activity while the class is in session.

Academic Accommodations

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to the instructor (or to a TA) as early in the semester as possible. DSP is located in STU 301 and is open 8.30 AM to 5.00 pm Monday through Friday. Website and contact information for DSP: `http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html` (213) 740–0776 (Phone), (213) 740–6948 (TDD only), (213) 740–8216 (FAX), ability@usc.edu

Statement on Academic Conduct and Support Systems

Academic Conduct

Plagiarism – presenting someone else's ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, "Behavior Violating University Standards" `policy.usc.edu/scampus--part--b`. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, `policy.usc.edu/scientific--misconduct`.

Support Systems

Student Health Counseling Services – (213) 740-7711 – 24/7 on call
`engemannshc.usc.edu/counseling`
Free and confidential mental health treatment for students, including short–term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline – 1 (800) 273-8255 – 24/7 on call
`suicidepreventionlifeline.org`
Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) – (213) 740–4900 – 24/7 on call
`engemannshc.usc.edu/rsvp`
Free and confidential therapy services, workshops, and training for situations related to gender–based harm.

Office of Equity and Diversity (OED) | Title IX – (213) 740-5086
`equity.usc.edu`, `titleix.usc.edu`
Information about how to get help or help a survivor of harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following protected characteristics: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations.

Bias Assessment Response and Support – (213) 740-2421
`studentaffairs.usc.edu/bias--assessment--response--support`
Avenue to report incidents of bias, hate crimes, and microaggressions for appropriate investigation and response.

The Office of Disability Services and Programs – (213) 740-0776
`dsp.usc.edu`
Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy – (213) 821-4710
`studentaffairs.usc.edu/ssa`
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC – (213) 740-2101
`diversity.usc.edu`
Information on events, programs and training, the Provost's Diversity and Inclusion Council,

Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency – UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call
`dps.usc.edu,emergency.usc.edu`
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety – UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call
`dps.usc.edu`
Non–emergency assistance or information.

Last updated: August 9, 2020