LPO 8852: Regression II

Sean P. Corcoran

Fall 2024

E-mail: sean.corcoran@vanderbilt.edu
Office: Wyatt Center 410N
Office Hours: Wednesdays 9-11 am (or by appt.)
https://calendly.com/sean-p-corcoran

Web: **seanpcorcoran.org**
Classroom: Peabody Admin 104
Class Hours: TTh 2:35-3:50pm
Phone: (615) 322-8021

Course Description

This course builds on the concepts and tools learned in Regression I (LPO 8851). The focus is on making causal inferences from observational (i.e., non-experimental) data through the use of matching, longitudinal (panel) data, difference-in-differences designs, synthetic control, instrumental variables, regression discontinuity, and other statistical techniques. While we will cover the theory related to these methods, the emphasis will be on their practical, hands-on application. Examples will come primarily from education research, although the skills taught in this course are broadly transferable across the social, behavioral, and health sciences.

Prerequisites

Students are expected to have successfully completed Regression I (LPO 8851) or the equivalent. Students are also expected to be proficient in Stata. If you have concerns about your prior preparation for this class, please see me immediately.

Books

I make frequent use of the following books (shorthand name in parentheses). The starred titles are highly recommended for this course.

★ (MM) Angrist, Joshua D., & Pischke, Jörn-Steffen. (2015). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press. See http://masteringmetrics.com/

★ (MIX) Cunningham, Scott. (2021). *Causal Inference: The Mixtape*. New Haven: Yale University Press. See https://mixtape.scunning.com/

★ (HK) Huntington-Klein, Nick. (2022). *The Effect: An Introduction to Research Design and Causality.* Boca Raton, FL: CRC Publishing. See https://theeffectbook.net/

★ (MW) Murnane, Richard J., & Willett, John B. (2011). *Methods Matter: Improving Causal Inference in Educational and Social Research*. New York: Oxford University Press.

(MHE) Angrist, Joshua D., & Pischke, Jörn-Steffen. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press. See https://www.mostlyharmlesseconometrics.com/

(C&T) Cameron, A. Colin, & Trivedi, Pravin K. (2022). *Microeconometrics Using Stata, 2nd Edition* (Volumes I & II). College Station, TX: Stata Press.

(G&T) Glewwe, Paul, & Todd, Petra. (2022). *Impact Evaluation in International Development: Theory, Methods, and Practice*. The World Bank Group. See https://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-1497-6

(WOOL) Wooldridge, Jeffrey M. (2016). *Introductory Econometrics: A Modern Approach, 6th Edition*. Boston: Cengage Publishing.

If you have recently taken LPO 8851, you should already have the Wooldridge text. Other readings on this syllabus and other resources will be made available via Github. You can also easily locate them using Google Scholar.

Course Structure

The class will meet twice weekly, in person. Class meetings will be a mix of lecture and in-class lab exercises. Supplemental material may be provided online.

This is a graduate course designed for students at the doctoral and advanced master's level. I expect that you are motivated by a desire to learn the course material. Please come to class having carefully reviewed any relevant book chapters and supplemental materials on the reading list. Attendance in class and participation in in-class lab exercises is extremely important. For your reference I will record most lectures via Zoom. See the course Brightspace page for recordings.

Stata

Stata is the statistical software used in this course. I recommend buying a short-term or perpetual license. Stata has recently moved to a subscription model called StataNow where users pay an annual fee to license a continually-updated version of Stata (the current version is 18). It is rather expensive, $360 per year for Stata BE (handles mid-size datasets) and $510 for Stata SE (handles large datasets). There are discounted rates for group purchases (2 or more). A perpetual license does not expire but also does not entitle you to upgrades. See the Stata website for details and make sure you get the best price for your intended usage: https://www.stata.com/order/educational-license-options/. I am currently using Stata 18, but other recent vintages are fine

for this class (e.g., 14-17). Be aware that small differences exist between versions, and that files created in recent versions of Stata may not open in older versions. If you don't wish to purchase Stata, it is freely available to you in campus computer labs. Vanderbilt also makes Stata available to you virtually through VMWare: https://anywherevu.vanderbilt.edu/portal/webclient/index.html.

There are many excellent resources for learning Stata. See this site for a useful starting point: https://www.stata.com/links/. There are some handy Stata "cheat sheets" posted here: https://geocenter.github.io/StataTraining/portfolio/01_resource/. I will also upload numerous other Stata references and links to Github.

Course Requirements

Your grade for the course will be based on approximately **nine** problem sets (50%), an in-class midterm (25%), and a final exam (25%). The problem sets will vary in length and points possible, but each will be weighted equally when calculating your final grade, using the percent correct on each. Unfortunately, due to the advent of AI tools, exams must be administered in person.

The (tentative) schedule for problem sets is shown in the schedule below. Due dates and the number of problem sets are subject to change based on the pace of the course. Please submit your problem set solutions via email to sean.corcoran@vanderbilt.edu. Include your last name and problem set number in the filename (e.g., *Corcoran_PS1.pdf*). Late assignments will not be accepted, particularly after solutions have been provided or discussed in class.

Unless otherwise indicated, the file you submit to me should be a log of your Stata session, saved as a text file (with the .txt extension) or—better yet—converted to a PDF. Begin by copying the problem set instructions into the Stata do-file editor. Comment out the questions. Insert after each question the commands you used to respond to that question. The resulting log file will include the instructions (in the form of comments), your commands, and the output. Edit this file as appropriate, for example by adding interpretations of your output and any other commentary that might be asked for. Graphical output can be submitted separately, preferably as a PDF file. (Combine all your PDFs if possible). Please clean up your submitted results so they are readable and look professional. You are encouraged to work together on the problem sets, but all work submitted must be that of the individual student. Duplicate assignments will not be accepted.

Based on feedback in previous semesters, I will answer questions about the problem sets in class but not devote much class time to the solutions. Detailed solutions will be posted on Github.

Other Important Information

1. **Github**: All materials pertaining to this course, including lecture notes, problem sets, and datasets, will be available on Github (https://github.com/spcorcor18/LPO-8852). Check in frequently for new material and announcements. Lectures will be posted in advance of class, but occasional delays and revisions are to be expected. The course is stored in what Github calls a "repository". You can "clone" (sync) this repository to your local drive

using Github Desktop (https://desktop.github.com/). I recommend this easy approach to staying up to date with all of the course materials.

2. **Classroom etiquette**: Please bring your laptop to class. To help promote a productive learning environment, please devote your time and attention to the class itself. Please do not use social media, text messaging, email, or other digital distractions while in class. Please silence your cell phone as well.

3. **Artificial intelligence**: Do not use Chat GPT or other artificial intelligence tools in this course, for coding or completing problem sets. The goal is for **you** to learn the concepts and coding skills taught in this course. I will not accept submitted assignments that have clearly relied on AI.

4. **Health and safety**: Our mutual commitment to health and safety is vital. Toward that end, all students are expected to adhere to Vanderbilt health and safety protocols. Guidance may be updated throughout the semester.

5. **Names and pronouns**: If you would like to use a different name or pronouns than those provided through YES, please let me know at any time prior to or during the semester. Information is available through the LGBTQI Life offices about how to change either or both of these in YES.

6. **Honor code**: All work submitted in this course is governed by provisions of the Vanderbilt University Honor Code, found in the student handbook: http://www.vanderbilt.edu/student_handbook/the-honor-system/. Any attempt to pass off someone else's work as your own is a violation of the Honor Code, and there are many ways this can happen beyond blatant cheating. If you have any doubts about how the Honor Code applies to your work in this class, please ask me for clarification. Uncertainty about application of the Honor Code does not excuse a violation.

7. **Classroom accommodations**: Vanderbilt is committed to equal opportunity for students with disabilities. If you need course accommodations due to a disability, please contact VU Student Access Services to initiate the process: https://www.vanderbilt.edu/student-access/. After SAS has notified me of relevant accommodations, we will discuss how these accommodations may best be approached in this class, and I will facilitate the accommodations.

8. **Mandatory reporter obligation**: All university faculty and administrators are mandatory reporters. What this means is that all faculty, including me, must report allegations of sexual misconduct and intimate partner violence to the Title IX Coordinator (615-343-9004). In addition, all faculty are obligated to report any allegations of discrimination. I am willing to discuss such incidents with you, but I can only do so in the context of us both understanding my reporting obligations. If you want to talk with someone in confidence, officials in the Student Health Center, the University Counseling Center, and the Office of the Chaplain and Religious Life (when acting as clergy) can maintain confidentiality. In addition, officials in the Project Safe Center (Crisis Hotline: 615-322-7233) have limited confidentiality, in that they have to report the incidents they are told of, but can do so without providing identifying information about the victim(s). The Project Safe Center

(https://www.vanderbilt.edu/projectsafe/) serves as the central resource for those impacted by sexual misconduct and intimate partner violence and can assist with navigating all facets of the University's resource and support network and other processes

9. **Mental health and wellness**: If you are experiencing undue stress during the semester that may be interfering with your ability to perform academically, Vanderbilt's Student Care Network offers a range of support services. I am available to speak with you about stresses related to your work in this course, and I can assist you in connecting with the Student Care Network. The Office of Student Care Coordination (OSCC) is the central and first point of contact to help you navigate and connect to appropriate resources. You can schedule an appointment with the OSCC at https://www.vanderbilt.edu/carecoordination/ or call 615-343-WELL. The Student Care Network also offers drop-in services on campus on a regular basis. You can find a calendar of services at https://www.vanderbilt.edu/studentcarenetwork/satellite-services/

   If you or someone you know needs to speak with a professional counselor immediately, the University Counseling Center offers Urgent Care Counseling. Students should call the UCC at (615) 322-2571 during office hours to speak with an urgent care clinician. You can also reach an on-call counselor after hours or on the weekends by calling (615) 322-2571 and pressing option 2 at any time. You can find additional information at https://www.vanderbilt.edu/ucc/.

Course outline

Starred readings are highly recommended. Others are for your reference.

Lecture 1: Potential outcomes and treatment effects

- ★ MIX, *Potential Outcomes Causal Model* (especially pp. 119-148)
- ★ HK chapter 10, *Treatment Effects*
- ★ MM chapters 1-2, *Randomized Trials* and *Regression*

  G&T chapters 3, 5-7

  MHE chapter 2, *The Experimental Ideal*

  MW chapter 4, *Investigator-Designed Randomized Experiments*

Lecture 2: Matching and weighting estimators

- ★ MIX, *Matching and Subclassification*
- ★ HK chapter 14, *Matching*
- ★ MM chapter 2, *Regression* (especially pp. 47-59)
- ★ Guo & Fraser (2015), *Propensity Score Analysis: Statistical Methods and Applications, 2e*.

  G&T chapter 13, *Matching Methods*

  MW chapter 12, *Dealing with Bias in Treatment Effects Estimated from Nonexperimental Data*

  Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1), 31–72. https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-6419.2007.00527.x

  Imbens, G. W. (2015). Matching Methods in Practice: Three Examples. *Journal of Human Resources*, 50(2), 373–419. https://doi.org/10.3368/jhr.50.2.373

  Morgan, S. L., & Harding, D. J. (2006). Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice. *Sociological Methods & Research*, 35(1), 3–60. https://doi.org/10.1177/0049124106289164

Lecture 3: Panel data

- ★ MIX, *Panel Data*
- ★ HK chapter 16, *Fixed Effects*

C&T chapters 8-9

WOOL chapters 13-14

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2022). When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics*, 138(1), 1-35. https://doi.org/10.1093/qje/qjac038

Cameron, C.A. & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

MW chapter 7, *Experimental Research When Participants Are Clustered within Intact Groups*

Lecture 4: Difference-in-differences

★ MIX, *Difference-in-Differences*

★ MM chapter 5, *Differences-in-Differences*

★ Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *Journal of Econometrics*, 235(2), 2218-2244. https://doi.org/10.1016/j.jeconom.2023.03.008

★ Wing, C., Yozwiak, M., Hollingsworth, A., Freedman, S., & Simon, K. (2024). Designing Difference-in-Difference Studies with Staggered Treatment Adoption: Key Concepts and Practical Guidelines. *Annual Review of Public Health*, 45, 485-505. https://doi.org/10.1146/annurev-publhealth-061022-050825

HK chapter 18, *Difference-in-Differences*

G&T chapters 11-12

MHE chapter 5, *Parallel Worlds: Fixed Effects, Difference-in-Differences, and Panel Data*

WOOL chapter 13

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1), 249–275. https://academic.oup.com/qje/article-abstract/119/1/249/1876068

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, 225(2), 254-277. https://doi.org/10.1016/j.jeconom.2021.03.014

Jakiela, P. (2021). Simple Diagnostics for Two-Way Fixed Effects. Working paper available at https://arxiv.org/pdf/2103.13229.pdf

Olden, A., & Møen, J. (2022). The Triple Difference Estimator. *The Econometrics Journal*. https://doi.org/10.1093/ectj/utac010

Lecture 5: Event studies

★ HK chapter 17, *Event Studies*

★ Clarke, D., & Tapia-Schythe, K. (2021). Implementing the Panel Event Study. *The Stata Journal*, 21(4), 853-884. https://journals.sagepub.com/doi/abs/10.1177/1536867X211063144

★ Miller, D. L. (2023). An Introductory Guide to Event Study Models. *Journal of Economic Perspectives*, 37(2), 203-230. https://www.aeaweb.org/articles?id=10.1257/jep.37.2.203

Lecture 6: Synthetic control methods

★ MIX, *Synthetic Control*

★ Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2), 391-425. https://www.aeaweb.org/articles?id=10.1257/jel.20191450

Bonander, C., Humphreys, D., & Degli Esposti, M. (2021). Synthetic Control Methods for the Evaluation of Single-Unit Interventions in Epidemiology: A Tutorial. *American Journal of Epidemiology*, 190(12), 2700-2711. https://doi.org/10.1093/aje/kwab211

Lecture 7: Instrumental variables

★ MM chapter 3, *Instrumental Variables*

★ HK chapter 9, *Instrumental Variables*

★ MIX, *Instrumental Variables*

C&T chapter 7

G&T chapter 15

MHE chapter 4, *Instrumental Variables in Action: Sometimes You Get What You Need*

MW chapters 10-11, *Introducing Instrumental Variables Estimation* and *Using IVE to Recover the Treatment Effect in a Quasi-Experiment*

WOOL chapter 15

Lecture 8: Regression discontinuity

★ MM chapter 4, *Regression Discontinuity Designs*

★ MIX, *Regression Discontinuity*

★ Bloom, H. S. (2012). Modern Regression Discontinuity Analysis. *Journal of Research on Educational Effectiveness*, 5(1), 43–82. https://doi.org/10.1080/19345747.2011.578707

★ Cattaneo, Idrobo and Titiunik. (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press. https://rdpackages.github.io/references/Cattaneo-Idrobo-Titiunik_2020_CUP.pdf

★ Cattaneo, Idrobo and Titiunik. (2021). *A Practical Introduction to Regression Discontinuity Designs: Extensions.* Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press. https://rdpackages.github.io/references/Cattaneo-Idrobo-Titiunik_2021_CUP.pdf

G&T chapter 14

MHE chapter 6, *Getting a Little Jumpy: Regression Discontinuity Designs*

MW chapter 9, *Estimating Causal Effects Using a Regression-Discontinuity Approach*

Imbens, G. W., & Lemieux, T. (2008). Regression Discontinuity Designs: A Guide to Practice. *Journal of Econometrics*, 142(2), 615–635. http://dx.doi.org/10.1016/j.jeconom.2007.05.001

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2), 281-355. https://doi.org/10.1257/jel.48.2.281

What Works Clearing House Standards Handbook Version 4.0 (2017), https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_standards_handbook_v4_draft.pdf

Other recommended readings: surveys

- Abadie, A., & Cattaneo, M. D. (2018). Econometric Methods for Program Evaluation. *Annual Review of Economics*, 10(1), 465–503. https://doi.org/10.1146/annurev-economics-080217-053402
- Angrist, J. (2022). Empirical Strategies in Economics: Illuminating the Path from Cause to Effect. National Bureau of Economic Research Working Paper Series, No. 29726. https://doi.org/10.3386/w29726
- Schlotter, M., Schwerdt, G., & Woessmann, L. (2011). Econometric Methods for Causal Evaluation of Education Policies and Practices: A Non-Technical Guide. *Education Economics*, 19(2), 109–137. https://doi.org/10.1080/09645292.2010.511821

Other recommended readings: effect size

- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, 2(3), 172–177. https://doi.org/10.1111/j.1750-8606.2008.00061.x
- Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*. https://doi.org/10.3102/0013189X20912798

Schedule at a glance

Aug 22	Lecture 1: Potential outcomes and treatment effects	PS1 assigned
Aug 27	Lecture 1: Potential outcomes and treatment effects	
Aug 29	Lecture 1: Potential outcomes and treatment effects	PS1 due
Sep 3	Lecture 2: Matching and weighting estimators	PS2 assigned
Sep 5	Lecture 2: Matching and weighting estimators	
Sep 10	Lecture 2: Matching and weighting estimators	PS2 due, PS3 assigned
Sep 12	Lecture 2: Matching and weighting estimators	
Sep 17	Lecture 3: Panel data	PS3 due, PS4 assigned
Sep 19	Lecture 3: Panel data	
Sep 24	Lecture 4: Difference-in-differences	PS4 due, PS5 assigned
Sep 26	Lecture 4: Difference-in-differences	
Oct 1	Lecture 4: Difference-in-differences	PS5 due
Oct 3	Review for midterm exam	
Oct 8	**Midterm exam**	
Oct 10	**NO CLASS - Fall break**	
Oct 15	Lecture 4: Difference-in-differences	PS6 assigned
Oct 17	Lecture 5: Event studies	
Oct 22	Lecture 5: Event studies	
Oct 24	Lecture 6: Synthetic control method	PS6 due, PS7 assigned
Oct 29	Lecture 6: Synthetic control method	
Oct 31	Lecture 7: Instrumental variables	PS7 due, PS8 assigned
Nov 5	Lecture 7: Instrumental variables	
Nov 7	Lecture 7: Instrumental variables	PS8 due
Nov 12	Lecture 8: Regression discontinuity	PS9 assigned
Nov 14	Lecture 8: Regression discontinuity	
Nov 19	Lecture 8: Regression discontinuity	
Nov 21	**NO CLASS - APPAM**	
Nov 26	**NO CLASS - Thanksgiving**	
Nov 28	**NO CLASS - Thanksgiving**	
Dec 3	Lecture 8: Regression discontinuity	PS9 due
Dec 5	Review for final	Last day of classes
Dec 7	**Final exam**	

Notes:

- Due dates and the number of problem sets are subject to change. If fewer than 9 problem sets are assigned, I will give equal weight to the problem sets assigned.
- I will be traveling on **Tuesday October 8** (midterm exam date) and **Thursday November 21** for APPAM.
- Many of you may be interested in attending the Zachary Bleemer seminar in the Economics Department on Monday, September 30 (Buttrick Hall Room 344)