LPO 8852: Regression II

Sean P. Corcoran

Fall 2021

E-mail: sean.corcoran@vanderbilt.edu
Office: Payne Hall 205C
Office Hours: Wednesdays 9-10:3am (or by appt.)
https://calendly.com/sean-p-corcoran

Web: **seanpcorcoran.org**
Classroom: Payne 108
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, 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 and applications will come primarily from education research, although the skills taught in this course are broadly transferable across subject areas in 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 moderately proficient in Stata. If you have concerns about your prior preparation for this class, please see me immediately.

Books

There are four recommended (not required) texts for this course. These are referenced in the reading list as MM, MIX, WOOL, and C&T:

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

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

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

- Cameron, A. Colin, & Trivedi, Pravin K. (2010). *Microeconometrics Using Stata, Revised Edition*. College Station, TX: Stata Press.

If you have recently taken LPO 8851, you should already have the Wooldridge text. Other books that I reference often include the following. You may find them helpful in your own collection. The first of these is referenced as MHE:

- Angrist, Joshua D., & Pischke, Jörn-Steffen. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

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

- Mitchell, Michael N. (2012). *A Visual Guide to Stata Graphics, 3rd Edition*. College Station, TX: Stata Press.

Other readings listed in the course schedule below will be made available via Github.

Course Structure

The class will meet twice weekly, in person. Class meetings will be a mix of lecture and in-class exercises. In some cases, supplemental material will be provided online.

This is a graduate course designed for students at the doctoral and advanced master's level. I expect that students enrolled in the course are motivated by a desire to learn the course material. Please come to class having carefully read any relevant textbook chapters and supplemental materials. Attendance in class and participation in in-class lab exercises is extremely important.

Stata

Stata is the statistical software used in this course. I recommend the most recent release (Version 17), but other recent vintages are fine (e.g., 13-16). I presently use Stata 15. Be aware that small differences exist between versions, and that files created in recent versions of Stata may not open in older versions. Purchasing options are available via the following website: https://www.stata.com/order/new/edu/gradplans/student-pricing/. A 6-month license for Stata/BE can be purchased for $48, but if you intend to use Stata in your own work I recommend purchasing a perpetual license of Stata/SE or the more powerful multi-processor (MP) version. Stata is freely available to you in the Wyatt 132 computer lab and elsewhere on campus. Vanderbilt also makes Stata available to you virtually through VMWare: https://anywherevu.vanderbilt.edu/portal/webclient/index.html.

There are many great resources for learning Stata, including the Acock text noted above. UCLA has some nice online resources for learning Stata (https://stats.idre.ucla.edu/stata/) and the Stata YouTube site is also quite helpful. There are some handy Stata "cheat sheets" posted here: https://geocenter.github.io/StataTraining/portfolio/01_resource/. I will also upload some Stata references to Github.

Course Requirements

Your grade for the course will be based on **ten** problem sets (50%), a midterm (25%), and 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.

The (tentative) schedule of problem set assignments is shown in the course schedule below. These are subject to change based on the pace of the course. Please submit your problem set solutions to me via email at 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 problem set 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. 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.

Other Important Information

1. **Github**: All materials pertaining to this course, including lecture notes, problem sets, and datasets, will be available on Vanderbilt's Github platform (https://github.com/spcorcor18/LPO-8852). Check in frequently for new materials and announcements. Most materials will be posted in advance of class, but occasional delays 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 Facebook/Instagram, text messaging, email, or other digital distractions while in class. Please silence your cell phone as well.

3. **COVID-19**: Vanderbilt is requiring masks indoors, except where physical distancing (6 feet) is possible. Instructors are not required to wear masks as long as they remain 6 feet away from others. I plan to not wear a mask when lecturing—assuming physical distancing is possible—but will have one on hand to wear as necessary. Quarantine is not required if you are vaccinated, have close contact with someone with COVID-19, and are asymptomatic. However, if you test positive for COVID-19, you will be required to isolate for 10 days. If this occurs, I will work with students who miss class on a case by case basis. As noted, all class materials—including lecture notes—will be available on Github. For the latest on Vanderbilt protocols for COVID-19, see: https://www.vanderbilt.edu/coronavirus/.

4. **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.

5. **Academic integrity**: All academic work at Vanderbilt is done under the Honor System. Students are expected to conform to the highest standards of academic integrity in this course. Any attempt to pass off someone else's work as your own is a violation of this standard, and there are many ways this can happen beyond blatant cheating. Full details of the Vanderbilt Honor System may be found here: `http://www.vanderbilt.edu/student_handbook/the-honor-system/` If you have any doubts about how the Honor Code applies to your work in this class, please ask me—not another student—for clarification. Uncertainty about application of the Honor Code does not excuse a violation.

6. **Accommodations**: Vanderbilt is committed to equal opportunity for students with disabilities, as am I. 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.

7. **Mandatory reporter obligation**: All university faculty and administrators are mandatory reports. What this means is that all faculty, including me, must report allegations of sexual misconduct and intimate partner violence to the Title IX Coordinator. In addition, all faculty are obligated to report any allegations of discrimination to the Title IX Coordinator (615-343-9004).

   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 officials in the Office of the Chaplain and Religious Life (when acting as clergy) can all 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).

8. **Mental health and wellness**: If you are experiencing undue personal and/or academic stress during the semester that may be interfering with your ability to perform academically, Vanderbilt's Student Care Network offers a range of services to assist and support you. 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 students navigate and connect to appropriate resources on and off-campus, develop a plan of action, and provide ongoing support. 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 Crisis Care Counseling during the summer and academic year. Students may come directly to the UCC and be seen by the clinician on call, or may call the UCC at (615) 322-2571 to speak with a clinician. You can find additional information at https://www.vanderbilt.edu//ucc/

Class schedule

The schedule is tentative and subject to change. Starred readings are required/highly recommended the associated lecture. Others are for your reference.

Lecture 1 (Aug 26): Introduction: regression and casuality

★ MM chapter 2 and its appendix

• 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

Lecture 2 (Aug 31 and Sept 2): Regression estimation and inference, bootstrapping (review)

★ MM chapter 2 and its appendix

★ MIX, *Probability and Regression Review* (especially pp. 36-93)

• WOOL chapters 2-6, 8

• C&T chapters 3 and 5

★ Stine, R. (1989). An Introduction to Bootstrap Methods: Examples and Ideas. *Sociological Methods & Research* 18(2-3): 243–291. https://doi.org/10.1177/0049124189018002003

• *Problem set 1 assigned Aug 31, due Sept 7*

Lecture 3 (Sept 7 and 9): Matching estimators I

★ MM chapter 2 (especially pp. 47-59)

★ MIX, *Matching and Subclassification*

★ 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

• 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

• Guo & Fraser (2015), *Propensity Score Analysis: Statistical Methods and Applications, 2e*.

• *Problem set 2 assigned Sept 7, due Sept 14.*

Lecture 4 (Sept 14 and 16): Matching estimators II

★ 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

• Wilde, E. T., & Hollister, R. (2007). How Close is Close Enough? Evaluating Propensity Score Matching Using Data from a Class Size Reduction Experiment. Journal of Policy Analysis and Management, 26(3), 455–477. https://doi.org/10.1002/pam.20262

• Guo & Fraser (2015)

• *Problem set 3 assigned Sept 14, due Sept 21.*

Lecture 5 (Sept 21 and 23): Difference-in-differences I

★ MM chapter 5

★ MIX, *Difference-in-Differences*

• MHE chapter 5

• WOOL chapter 13

• Baicker, K. and T. Svoronos (2019). Testing the Validity of the Single Interrupted Time Series Design. National Bureau of Economic Research Working Paper No. 26080. http://www.nber.org/papers/w26080

• Hallberg, K., Williams, R., Swanlund, A., & Eno, J. (2018). Short Comparative Interrupted Time Series Using Aggregate School-Level Data in Education Research. *Educational Researcher*, 47(5), 295–306. https://doi.org/10.3102/0013189X18769302

• *Problem set 4 assigned Sept 21, due Sept 28.*

Lecture 6 (Sept 28 and 30): Difference-in-differences II

★ Goodman-Bacon, A. (forthcoming). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*.

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

• 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

MIDTERM (Oct 7)

• Review for the midterm: Oct 5.

Lecture 7 (Oct 19 and 21): Panel data I

★ C&T chapter 8

★ WOOL chapter 13

★ MIX, *Panel Data*

• *Problem set 5 assigned Oct 19, due Oct 26.*

Lecture 8 (Oct 26 and 28): Panel data II

★ C&T chapter 8

★ WOOL chapter 14

• Raudenbush, S. W. (2009). Adaptive Centering with Random Effects: An Alternative to the Fixed Effects Model for Studying Time-Varying Treatments in School Settings. *Education Finance and Policy*, 4(4), 468–491. https://doi.org/10.1162/edfp.2009.4.4.468

• *Problem set 6 assigned Oct 26, due Nov 2*

Lecture 9 (Nov 2 and 4): Instrumental variables I

★ MM chapter 3

★ C&T chapter 6

★ MIX, *Instrumental Variables*

• WOOL chapter 15

• *Problem set 7 assigned Nov 2, due Nov 9.*

Lecture 10 (Nov 9 and 16): Instrumental variables II

★ MHE chapter 4

★ C&T chapter 6

★ MIX, *Instrumental Variables*

• *Problem set 8 assigned Nov 9, due Nov 16.*

Lecture 11 (Nov 18 and 30): Regression discontinuity

★ MM chapter 4

★ 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

- 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

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

- *Problem set 9 assigned Nov 18, due Dec 2.*

THANKSGIVING BREAK (Nov 23 and 25)

- *No class—enjoy!*

Lecture 12 (Dec 2 and 7): Clustered data and inference

- ★ Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When Should You Adjust Standard Errors for Clustering? National Bureau of Economic Research Working Paper No. 24003. http://www.nber.org/papers/w24003

- ★ Murnane & Willett, chapter 7, "Experimental Research When Participants Are Clustered within Intact Groups"

- Colin Cameron, 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

- C&T chapter 13

- *Problem set 10 assigned Dec 2, due Dec 9.*

FINAL EXAM (Dec 7-12)

- Final exam will be a take-home exam.

Schedule at a glance

Aug 26	Lecture 1: Introduction: regression and causality	
Aug 31	Lecture 2: Regression estimation and inference (review)	PS1 assigned
Sep 2	Lecture 2: Regression estimation and inference (review)	
Sep 7	Lecture 3: Matching estimators (I)	PS2 assigned
Sep 9	Lecture 3: Matching estimators (I)	
Sep 14	Lecture 4: Matching estimators (II)	PS3 assigned
Sep 16	Lecture 4: Matching estimators (II)	
Sep 21	Lecture 5: Difference-in-differences (I)	PS4 assigned
Sep 23	Lecture 5: Difference-in-differences (I)	
Sep 28	Lecture 6: Difference-in-differences (II)	
Sep 30	Lecture 6: Difference-in-differences (II)	
Oct 5	Review for midterm	
Oct 7	**Midterm**	
Oct 12	**NO CLASS - Prof. Corcoran traveling**	
Oct 14	**NO CLASS - VU Fall break**	
Oct 19	Lecture 7: Panel data (I)	PS5 assigned
Oct 21	Lecture 7: Panel data (I)	
Oct 26	Lecture 8: Panel data (II)	PS6 assigned
Oct 28	Lecture 8: Panel data (II)	
Nov 2	Lecture 9: Instrumental variables (I)	PS7 assigned
Nov 4	Lecture 9: Instrumental variables (I)	
Nov 9	Lecture 10: Instrumental variables (II)	PS8 assigned
Nov 11	**NO CLASS - APPAM**	
Nov 16	Lecture 10: Instrumental variables (II)	
Nov 18	Lecture 11: Regression discontinuity	PS9 assigned
Nov 23	**NO CLASS - Thanksgiving**	
Nov 25	**NO CLASS - Thanksgiving**	
Nov 30	Lecture 11: Regression discontinuity	
Dec 2	Lecture 12: Clustered data and inference	PS10 assigned
Dec 7	Lecture 12: Clustered data and inference	
Dec 7-12	**Final**	