

# LPO 7870-01: RESEARCH DESIGN AND DATA ANALYSIS II

Sean P. Corcoran

Spring 2024

E-mail: [sean.corcoran@vanderbilt.edu](mailto:sean.corcoran@vanderbilt.edu)

Office: Payne Hall 205C

Office Hours: Wednesdays 9-11 am (or by appt.)

<https://calendly.com/sean-p-corcoran>

Web: [seanpcorcoran.org](http://seanpcorcoran.org)

Classroom: Wyatt Center 132

Class: Monday 4:15-7:05 pm

Phone: (615) 322-8021

---

## Course Description

The central objective of this course is to instruct students in “state of the art” methods for the core tasks involved in research design and policy and program evaluation and to provide an understanding of when and how these methods can be most usefully applied to produce knowledge and evidence of program and policy effectiveness. Topics that will be covered include research design, causal attribution and validity, basic descriptive and inferential statistics, statistical power, multiple regression, and experimental and quasi-experimental methods (e.g., regression discontinuity and panel methods). The course will be taught using the Stata statistical software package and datasets commonly used in education policy research.

## Prerequisites

LPO 7860: Research Design and Data Analysis I or instructor consent.

## Books

I make frequent use of the following books (shorthand name in parentheses). The starred title is required. The others are highly recommended supplements to the main text.

★ (S&W) Stock, James H. and Watson, Mark W. (2020). *Introduction to Econometrics*. This text is currently in its 4th edition, but you are welcome to use the 2nd, 3rd, or 4th edition.

(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/>

(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>

(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.

If you have no or limited experience with Stata, you may benefit from the following simple introductory text (not required). I will provide other resources for learning Stata.

*A Gentle Introduction to Stata, Revised 6th edition* by Alan C. Acock, Stata Press, 2023.

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

## Course Structure

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

Please note this is a graduate course designed for students at the doctoral and 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 reviewed the required and recommended readings. Attendance and participation in class are extremely important given our one-per-week schedule.

## Stata

Stata is the statistical software used in this course. I recommend the most recent release (Version 18), but other recent vintages are fine (e.g., 14-17). 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 campus computer labs. Vanderbilt also makes Stata available to you virtually. For instructions see: <https://libanswers.library.vanderbilt.edu/faq/290980>.

There are many great 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/](https://geocenter.github.io/StataTraining/portfolio/01_resource/). I will also upload numerous other Stata references to Github.

## Course Requirements

Your grade for the course will be based on 5 problem sets (10% each for a total of 50%), a midterm exam (20%), a project (25%), and class participation (5%).

*Problem sets.* 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. These are intended to give you the opportunity to apply statistical concepts and tools learned in class, while also honing your programming skills. The (tentative) schedule of assignments is provided in the course schedule below. These are subject to change based on the pace of the course. Late assignments will not be accepted, particularly after problem set solutions have been provided or discussed in class. 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.

*Project.* We will provide you with a dataset and ask you to estimate the effects of an intervention using different methods of analysis. You can choose to work individually or in groups (max 3 persons per group). The project is **due Friday, April 26**.

*Participation.* Participation will be measured by class attendance and participation. You are expected to regularly attend class meetings and come prepared to fully participate in class discussions. Full participation in group projects will also count toward your participation score.

We expect final letter grades to be assigned according to the following distributional cut points: A (100-93); A- (92-90); B+ (89-87); B (86-80); B- (79-77); C+ (76-74); C (73-66); C- (65-63); D (62-55); F (<55). However, we may make modifications based on the observed distribution of scores at the end of the course.

## Teaching Assistants

The teaching assistants for this course are Sara White ([sara.m.white@vanderbilt.edu](mailto:sara.m.white@vanderbilt.edu)) and Kathryn Enriquez ([kathryn.enriquez@vanderbilt.edu](mailto:kathryn.enriquez@vanderbilt.edu)). The TAs will provide periodic homework help and review sessions. The first will be a review of Stata commands and procedures during the second week. Sara and Kathryn will administer a survey in the first week of class to determine optimal times for these sessions. A schedule 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-7870>). 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:** 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. **Chat GPT:** Do not use Chat GPT or other artificial intelligence tools in this course, for any purpose. 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. **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/](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.
7. **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 students 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/>.

## Class schedule

### Lecture 1 (Jan 11): Research design for causal inference

- ★ S&W chapter 1, *Economic Questions and Data*
- ★ M&W chapters 1-3, *The Importance of Theory and Designing Research to Address Causal Questions*
- G&T chapters 1-5, *The Basics of Impact Evaluation*

### NO CLASS (Jan 15): MLK, Jr. Holiday

### Lecture 2 (Jan 22): Review of descriptive and inferential statistics

- ★ S&W chapter 3, *Review of Statistics*
- Bradley, W. J. & Schaefer, K. C. (1998). "[Limitations of measurement in the social sciences](#)" in *The Uses and Misuses of Data and Models*, Sage Publications.
- Problem set 1 assigned

### Lecture 3 (Jan 29, Feb 5 and 12): Multiple regression fundamentals

- ★ S&W chapter 4, *Linear Regression with One Regressor*
- ★ S&W chapter 5, *Regression with a Single Regressor: Hypothesis Tests and Confidence Intervals*
- ★ S&W chapter 6, *Linear Regression with Multiple Regressors*
- ★ S&W chapter 7, *Hypothesis Tests and Confidence Intervals in Multiple Regression*
- Magnuson, K. A., Meyers, M. K., Ruhm, C. J., & Waldfogel, J. 2004. [Inequality in preschool education and school readiness](#). *American Educational Research Journal*, 41: 115–157.
- Gershenson, S., & Holt, S. B. 2015. [Gender Gaps in High School Students' Homework Time](#). *Educational Researcher*, 44(8), 432–441.
- Reber, S., & Smith, E. 2023. [College Enrollment Disparities: Understanding the Role of Academic Preparation](#). Washington, D.C.: The Brookings Institution.
- Problem set 1 due (Feb 5)
- Problem set 2 assigned (Feb 5)

**Lecture 4 (Feb 19): Statistical power**

- ★ M&W chapter 6, *Statistical Power and Sample Size*

Glennster, R. & Takavarasha, K. 2013. "Statistical Power" in *Running Randomized Evaluations: A Practical Guide*. Princeton University Press.

Cohen, J. 1992. A power primer. *Psychological Bulletin* 112-1: 155–159.

Problem set 2 due

Problem set 3 assigned

**Lecture 5 (Feb 26 and Mar 18): Nonlinear models and limited dependent variables**

- ★ S&W chapter 8, *Nonlinear Regression Functions*
- ★ S&W chapter 11, *Regression with a Binary Dependent Variable*

Orme, J. G. & Buehler, C. 2001. [Introduction to multiple regression for categorical and limited dependent variables](#). *Social Work Research*, 25(1): 49–61.

Kasman, M., & Loeb, S. 2013. [Principals' Perceptions of Competition for Students in Milwaukee Schools](#). *Education Finance and Policy*, 8(1), 43–73.

Ladd, H. F. 2011. [Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement?](#) *Educational Evaluation and Policy Analysis*, 33(2): 235–261.

Problem set 4 assigned (Mar 18)

**MIDTERM EXAM (Mar 4)—in class**

Problem set 3 due

**NO CLASS (Mar 11)—SPRING BREAK****Lecture 6 (Mar 25): Experimental and quasi-experimental methods**

- ★ S&W chapter 13, *Experiments and Quasi-Experiments*

Ravallion, M. 2001. [The mystery of the vanishing benefits: An introduction to impact evaluation](#). *World Bank Economic Review* 15(1): 115–140.

Glazerman, S. Mayer, D. & Decker, P. 2006. [Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes](#). *Journal of Policy Analysis and Management*, 25(1): 75–96.

## Lecture 7 (Apr 1): Regression discontinuity designs

- ★ 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>
- Dee, T. S. & Wyckoff, J. 2015. [Incentives, selection, and teacher performance: Evidence from IMPACT](#). *Journal of Policy Analysis and Management*, 34: 267-297.

Problem set 4 due

Problem set 5 assigned

## Lecture 8 (Apr 8): Time series and interrupted time series designs

- ★ S&W chapter 6, *Introduction to Time Series Regression and Forecasting*
- Gamoran, A. & An, B. P. 2016. [Effects of school segregation and school resources in a changing policy context](#). *Educational Evaluation and Policy Analysis* 38(1): 43–64.

## Lecture 9 (Apr 15 and 22): Panel data methods

- ★ S&W chapter 10, *Regression with Panel Data*
- Hillman, N. W., Tandberg, D. A., & Fryar, A. H. 2015. [Evaluating the impacts of “new” performance funding in higher education](#). *Educational Evaluation and Policy Analysis*, 37(4): 501–519.
- Jack, R., Halloran, C., Okun, J., & Oster, E. 2023. [Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts](#). *American Economic Review: Insights*, 5(2), 173-190.
- Jepsen, C., Troske, K., & Coomes, P. 2014. [The labor-market returns to community college degrees, diplomas, and certificates](#). *Journal of Labor Economics*, 32(1).

Problem set 5 due (Apr 15)

## PROJECT DUE (Apr 26)



## Schedule at a glance

---

---

Jan 11	Lecture 1: Research design for causal inference	
Jan 15	<b>NO CLASS - MLK, Jr. Holiday</b>	
Jan 22	Lecture 2: Review of descriptive and inferential statistics	PS1 assigned
Jan 29	Lecture 3: Multiple regression fundamentals	
Feb 5	Lecture 3: Multiple regression fundamentals	PS1 due
Feb 12	Lecture 3: Multiple regression fundamentals	PS2 assigned
Feb 19	Lecture 4: Statistical power	PS2 due
Feb 26	Lecture 5: Nonlinear models and limited dependent variables	PS3 assigned
Mar 4	<b>MIDTERM EXAM</b>	PS3 due
Mar 11	<b>NO CLASS - Spring Break</b>	
Mar 18	Lecture 5: Nonlinear models and limited dependent variables	PS4 assigned
Mar 25	Lecture 6: Experimental and quasi-experimental methods	
Apr 1	Lecture 7: Regression discontinuity designs	PS4 due
Apr 8	Lecture 8: Time series and interrupted time series designs	PS5 assigned
Apr 15	Lecture 9: Panel data methods	PS5 due
Apr 22	Lecture 9: Panel data methods	
Apr 26	<b>PROJECT DUE</b>	

---

---

Note: This schedule is tentative and subject to change, depending on the pace of the class.