

POLI 171 - Making Policy with Data

Spring 2020

04/03/20

Instructor: Duy Trinh

Email: ddtrinh@ucsd.edu

Office: <https://ucsd.zoom.us/my/ddtrinh>

Office Hours: Tuesday, 8:00-10:00am

TA: Nathan Mariano

Email: nathanmariano@ucsd.edu

Office: <https://ucsd.zoom.us/my/nmariano>

Office Hours: Wednesday 9:00-10:00am

1 Course Description

This course explores how we can make policy recommendations using data. Its overall goal is to provide a survey of the most commonly-used empirical tools for political science and public policy research. Our focus is design-based causal inference, or the use of statistical methods to answer research questions that concern the impact of some cause (e.g., an intervention, a change in institutions, passage of a law, changes in economic conditions, or policies) on a certain outcome (e.g., vote choice, income, election results, levels of violence, political attitudes). We cover a variety of causal inference designs and methods, including experiments, regression, matching, and difference-in-differences. We will analyze the strengths and weaknesses of these methods using applications from the real world.

The objectives of this course include:

1. Introducing an analytical framework of policy evaluation
2. Surveying the most commonly-used research designs for policy making
3. Introducing the most basic (and some of the most important) statistical concepts
4. Providing basic data analytical skills crucial for today's job market and academic research, including basic R programming

2 How to “attend” POLI 171

Due to the extraordinary circumstances that COVID-19 has created, this course will be conducted online. There will be a total of 20 lectures available as pre-recorded podcasts on

Canvas, as well as 9 live R lab sessions. Everyone is expected to “attend” all lectures and lab sessions.

2.1 Monday and Wednesday lectures

- Complete the assigned readings *before* lectures
- During our normal lecture hours (3:00pm - 3:50pm), log into Canvas, download the latest podcast and view it. Please do so at your pace, but latest before midnight (San Diego time)
- After you’ve finished viewing the podcast, log into Piazza and ask/discuss questions related to the lecture. Both the TA and the instructor will regularly check the board and answer questions posted, although everyone else is also encouraged to contribute to the discussion.
 - We also encourage you to ask questions on Piazza for clarifications, questions about concepts, problem sets, etc. You can sign up to the Piazza course page directly from this address: <https://piazza.com/ucsd/spring2020/poli171>. A student’s respectful and constructive participation on the forum will count toward his/her class participation grade. *Do not email your questions directly to the instructors or TAs* (unless they are of personal nature) —we will not be answering your questions regarding course materials or problem sets through email.

2.2 Friday labs

- Make sure you have Zoom installed on your device. For information about setting up and accessing Zoom: <https://blink.ucsd.edu/technology/file-sharing/zoom/index.html>.
- Before the lab’s meeting time, go to Canvas and download the latest assignment instruction under the module “Lab assignment instructions”
- On Friday 3:00pm, log into Zoom and go to the following meeting room: <https://ucsd.zoom.us/j/394492056>. The link is also available on Canvas. For security and privacy reasons, you will need a password in order to join the meeting room. We will distribute the password via class announcement.
- During class, collaborate among yourself and seek assistance from your TA to complete the assignment.
- After class ends, upload your completed assignment on Canvas under the module “Lab assignment”
 - If you are unable to complete the assignment during the lab’s allotted time, you may continue working on it after class. However, the assignment must be completed and submitted by midnight Friday (San Diego time). Late submissions will count against your participation grade

2.3 Office hours

- We expect you to rely mostly on Piazza for substantive questions. However, should you need individualized assistance (e.g. with your homework, problem sets, personal issues, etc.), both the instructor and the TA are more than happy to talk to you via Zoom
- To see us during office hours, first go to our respective [appoint.ly](#) sign-up link and register for a meeting slot. Please do so at least 1 hour before your desired meeting time.
 - Duy Trinh: <https://appoint.ly/s/ddtrinh/poli171>
 - Nathan Mariano: <https://appoint.ly/s/nmariano/15-min>
- Then, log into Zoom and go to our respective Zoom meeting rooms at your registered timeslot.
 - Duy Trinh: <https://ucsd.zoom.us/my/ddtrinh>
 - Nathan Mariano: <https://ucsd.zoom.us/my/nmariano>

2.4 Problem sets

- There will be four problem sets in total. Posted on Canvas one week before the due date; due at the beginning of lecture. All homeworks must be submitted on Canvas.
- HW1 will consist of basic R programming exercises, and HW2 will cover the potential outcomes framework and simple randomized experiments. You should work independently on both assignments; **do not collaborate with any other students on these assignments.**
- HW3 and HW4 will consist of slightly more complicated data analysis problems in R. **You may collaborate with a single other student on these assignments**, or if you prefer, you may work independently. If you collaborate, each of the coauthors should turn in a single, identical document, and you will each receive the same grade for the assignment.
 - We expect that you and your coauthor will collaborate on every portion of the assignment. You should understand (and be prepared to explain) every part of the document that you turn in.
- Late submission will be penalized (a day = 2/15 points; in other words, if your submission is late for 8 days, you will receive 0 points for that one).

3 Evaluation

- Four Problem Sets (15% each; 60% total).
- Final exam (25%). June 10 3:00-4:30 PM. Open book. To be proctored on Canvas.
- Class participation and lab assignments (15%).
- Optional Extra Credit Opportunity (5%). Posted May 22; due June 7 by midnight PST.

4 Books

4.1 Required Textbooks

1. Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, Christel M. J. Vermeersch. *Impact Evaluation in Practice*. The World Bank. Available on Canvas.
2. Imai, Kosuke. *Quantitative Social Science: An Introduction*, by Kosuke Imai, Princeton University Press. Selected chapters available on Canvas.

4.2 Optional Textbooks

1. Peng, Roger D. *R Programming for Data Science*. A Leanpub Book. For sale at <http://leanpub.com/rprogramming>.
2. Angrist, Joshua D. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. Available as an ebook through Roger.

5 Computation

The labs and problem sets of this course will use R, an open-source computing language that is very widely used in statistics and the social sciences.

- You can download R for free at <http://archive.linux.duke.edu/cran/>.
- RStudio is an integrated development environment designed for R. It is possible to program in R without using RStudio, but RStudio makes R programming much easier, especially for beginners. You can download RStudio for free at <https://www.rstudio.com/products/rstudio/download/>. Choose the RStudio Desktop Open Source License option, and select the installer designed for your platform.
- A nice way to get started is with the two video tutorials provided by Dan Goldstein: <http://www.dangoldstein.com/flash/Rtutorial1/Rtutorial1.html> and <http://www.dangoldstein.com/flash/Rtutorial2/Rtutorial2.html>

- Another good resource is the set of tutorials provided by DataCamp: <https://www.datacamp.com/>
- We will cover the basic tools that you need for the homework assignments in lab sessions.

6 Expectations and Policies

- If you seek a re-grade, you must email the TA within 24 hours of the assignment being returned to the class, and explain—in that email and in detail—why you believe you deserve reconsideration. The TA then has the ability to review the entire assignment, and he might increase your grade, decrease your grade, or keep the grade unchanged.
- This course covers graduate-level concepts with undergraduate-level math. The focus is on the intuition, and the math will not be particularly difficult. If you passed POLI 30, you should have no problem doing well in this course.
- Prior experience with R programming will be useful, but not required. If you haven't used R before, you should expect a steep learning curve. Although we will cover the basics during our four lab sessions, most of you will master R programming only through doing it yourselves and by learning from each other.

7 Course Schedule

Introduction

Reading: Gertler, Chapter 1, pp. 3-9; Imai, Chapter 1.

- March 30: Introduction and Course Overview
- April 01: R Basics
- April 03: **R Lab 1**: Zoom and R Basics

Module I: Causality and Potential Outcomes

Reading: Gertler, Chapter 3, pp 33-40; Imai, Chapter 2, pp 40-50

- April 6: The Potential Outcomes Framework
- April 8: The Potential Outcomes Framework
- April 10: **R Lab 2**: Subsetting

Reading: Gertler, Chapter 3, pp 40-47, Imai, Chapter 2, pp 59-65

- April 13: Omitted Variable Bias. Problem Set 1 assigned
- April 15: Selection into Treatment.
- April 17: **R Lab 3**: Omitted Variable Bias

Module II: Randomized Control Trials (RCT)

Reading: Gertler, Chapter 4, pp. 49-69; Imai Chapter 2, pp 53-58

- April 20: Randomization and Experiments. Problem Set 1 due
- April 22: Randomization and Experiments.
- April 24: **R Lab 4**: Randomization

Reading: Gertler, Chapter 11, pp 171-195

- April 27: Inference and Experiments.
- April 29: Inference and Experiments. Problem Set 2 assigned
- May 1: **R Lab 5**: Inference

Reading: Olken (2007) "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." Available on Canvas.

- May 4: Interaction Effects
- May 6: Heterogeneous Treatment Effects. Problem Set 2 due
- May 8: **R Lab 6**: RCT 3

Module III: Observational Studies

Reading: Gertler, Chapter 7

- May 11: Selection on Observables
- May 13: Matching. Problem Set 3 assigned
- May 15: **R Lab 7**: Matching

Reading: Imai, Chapter 4, pp 150-161

- May 18: Regression.
- May 20: Fixed Effects. Problem Set 3 due

- May 22: **R Lab 8**: Regression. Optional project assigned

Reading: Gertler, Chapter 6; Imai, Chapter 2, pp 65-68

- May 25: Difference in Differences. Problem Set 4 assigned
- May 27: Panel Data
- May 29: **R Lab 9**: Panel Data

Review and Final Exam

- June 1: Optional topics. Problem Set 4 due
- June 3: Summary and Conclusion
- June 5: Review Session
- June 7: Optional project due
- June 10: Final Exam. 3:00-4:30 PM. Open Book.