POLI 171: Making Policy with Data

Winter, 2018

Instructor: Cameron Sells

- *Email:* csells@ucsd.edu
- *Office:* SSB 352
- *Office Hours*: Tuesday, 12:00-1:50pm

Teaching Assistant: Leo Yang

- *Email:* yay103@ucsd.edu
- *Office:* SSB 351
- *Office Hours*: Wednesday, 9:00-10:50am

Course Description

This course explores how we can make policy recommendations using data. The overall goal of the course 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, difference-in-differences, and regression discontinuity designs. 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

This course meets twice a week for the ten weeks of the quarter. Sixteen of the class days will be devoted to lecture, and the other four will be R lab sessions that will take place in our normal classroom. Each meeting will begin with a policy brief that will demonstrate how the methods covered in this course can be used to evaluate real-world policies. You are expected to attend all lectures and lab sessions, and you are strongly encouraged to bring your laptop (with R installed) to the four lab sessions.

Evaluation

- Five Problem Sets (14% each; 70% total). Posted on TritonEd one week before the due date; due at the beginning of lecture. All homeworks must be typed and submitted as a hardcopy.
  - HW1 will consist of basic R programming exercises, and HW2 will cover the potential outcomes framework. You should work independently on both assignments; **do not collaborate with any other students on these assignments.**
  - HW3, HW4 and HW5 will consist of 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, you and your coauthor will turn in a single document, and you will each receive the same grade for the assignment.
- Final Paper (25%). Due Finals Week, Thursday, March 22, by 6pm in SSB 352.
- Participation and Attendance (5%).
- Optional Extra Credit Opportunity (5%). Posted March 1; due March 15.

Expectations and Policies

- Late assignments will not be accepted unless cleared with the instructor 24 hours before the due date.
- 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 has the authority to increase your grade, decrease your grade, or keep the grade unchanged.
- We expect you to attend all lectures and labs.
- There is no prerequisite for this course. However, POLI 30 or the equivalent is strongly recommended.
- 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 is not necessary, but 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.
- If you choose to collaborate with another student on HW3, HW4 or HW5, 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.
- All lecture slides and assignments will be available on this course's TritonEd page.

Books

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 here.

2. Peng, Roger D. *R Programming for Data Science.* A Leanpub Book. For sale here.

Optional Textbooks

1. Angrist, Joshua D. and Jorn-Steffen Pischke. *Mastering Metrics: The Path from Cause to Effect.* Princeton University Press.

2. Angrist, Joshua D. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press. Available as an ebook through Roger.

Computation

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

- R runs on a wide variety of UNIX platforms, Windows and MacOS. R makes programming very easy, has strong graphical capabilities, and also contains canned functions for most commonly-used estimators.

- You can download R for free here. Select the link that matches your machine's platform. If you are using Windows, click on the "install R for the first time" link, and then click on the link for the latest version of R. If you are using MacOS, scroll down and find the version of R that matches your version of MacOS (Unless you are running an old version of OS X, you probably want "R-3.4.3.pkg," the latest version of R. You can check your version of MacOS by clicking on the apple icon and then on the "About This Mac" option.).

- 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 here. 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: Tutorial 1 and Tutorial 2

- Another good resource is the set of tutorials provided by DataCamp

- The web provides many other great tutorials and resources for learning R.

- We will cover the basic tools that you need for the homework assignments in lab sessions held in our normal classroom on January 11, February 1, February 20, and March 6.

Course Schedule

Introduction

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

- January 9: Introduction and Course Overview
- January 11: **R Lab 1**: R Basics

Unit I: Causality and Potential Outcomes

Reading: Gertler, Chapter 3

- January 16: The Potential Outcomes Framework
- January 18: Omitted Variable Bias and Selection Into Treatment
  - **HW1 (Computation in R) Due**

Unit II: Experiments

Reading: Gertler, Chapter 4, pp. 49-69

- January 23: Randomization and Experiments
- January 25: Inference and Experiments
  - **HW2 (Potential Outcomes) Due**
- January 30: What Can Go Wrong? Attrition, Noncompliance, and Spillovers
- February 1: **R Lab 2**: Experiments
- February 6: Clustering, Blocking and Multiple Treatment Arms

Unit III: Basic Observational Studies

Reading: Gertler, Chapter 7

- February 8: Selection on Observables
  - **HW3 (Experiments) Due**
- February 13: Regression
- February 15: Matching
- February 20: **R Lab 3**: Regression and Matching

Unit IV: Advanced Observational Studies

Reading: Gertler, Chapters 6 and 5

- February 22: Difference-in-Differences
- February 27: Synthetic Control Methods
  - **HW4 (OLS and Matching) Due**
- March 1: Regression Discontinuity
  - **Extra Credit Assignment Announced**
- March 6: **R Lab 4**: Difference-in-Differences and Regression Discontinuity

Unit V: Instrumental Variables

Reading: Gertler, Chapter 4, pp. 69-79

- March 8: Instrumental Variables Theory and Critiques
- March 13: Fuzzy Regression Discontinuity and Encouragement Experiments
  - **HW5 (DiD and RD) Due**

Conclusion

- March 15: Summary and Conclusion
  - **Extra Credit Assignment Due**

Final paper due March 22 by 6pm in SSB 352