

---

# POLI 171: Making Policy with Data

Summer Session I - 2025

**Instructor:** Alex Richard Zhao

**Office Hours:** [Calendly](#) (Th 10:00AM-12:00PM)

**Canvas:** [Course Link](#)

**Email:** [arzhao@ucsd.edu](mailto:arzhao@ucsd.edu)

**Class Hours:** M/W 11:00AM-1:50PM

**Class Room:** [Center Hall 205](#)

---

## COURSE DESCRIPTION

This course explores how we can make policy recommendations using data. Its 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., election results, levels of violence, political attitudes, or development). We will introduce and cover a variety of causal identification strategies including experiments, regression, matching, and panel methods. We will analyze the strength and weaknesses of these methods using applications from the real world. In addition, we will practice implementing these methods throughout the course.

The objectives of this course include:

1. Understand Foundational Statistical Concepts: Students will be able to explain foundational statistical concepts for empirical social science research
2. Explain Causal Inference Strategies: Students will be able to explain the most commonly-used research designs for policy making and causal claims
3. Analyze and Critique Research Designs: Students will be able to analyze the assumptions needed for an empirical research design to provide policy recommendations
4. Develop Their Programming: Students will develop the fundamental data skills needed for today's job market and academic research, particularly programming in R
5. Draft A Research Proposal: Students will be able to design and draft a research proposal intended to answer a question of their interest with a causal identification strategy

## PREREQUISITES

POLI 5 or 5D or ECON 5 or 5D, and POLI 30 or 30D.

## MATERIALS

We will primarily utilize readings from [Quantitative Social Science](#) by Imai and [Impact Evaluation in Practice](#) by Gerter et. al., which will be made available on Canvas. Please do your best to read before class. Any other research articles will also be posted on Canvas.

---

We will use R, an open-source computing language widely used in statistics and the social sciences. R is free and can be downloaded directly from [this link](#). RStudio is an integrated development environment designed for R. It is possible to program in R without using RStudio, but RStudio makes working with R much easier. RStudio is also free and can be downloaded directly from [this link](#). Instructions for using R can be found in the *Hands-On Programming with R* guide. I encourage you to read through this guide to begin familiarizing yourself with R. We will cover the basic tools and foundations so you can succeed in the homework and lab sessions. If you do not have access to a laptop, I recommend reaching out to the campus [Laptop Lending Program](#)

## COURSE STRUCTURE

There are no in-class midterms or finals for this course. Your course grade will be based on a combination of attendance, four problem sets, a final research proposal, and extra credit opportunities. Any assignment involving R must be submitted with accompanying code. All results should be presentable so they can be easily understood. Use comments, pseudocode, and create subsections to make code more legible.

### Problem Sets

This class will include 4 different weekly problem sets due throughout the course and a final project. Again, you will also be graded based on attendance and will have the opportunity to submit a written assignment for extra credit. Elements of the final research proposal will be included as part of each problem set where the instructional staff will provide feedback and advice.

*Due Date:* The problem sets will be released every Monday at 8:00AM and will be due by Monday of the following week at 11:59PM unless otherwise noted.

*Collaboration:* Group work is allowed but students must turn in all of their work individually. For example, students should not submit another student's answers or computer code as their own. Students must write their full name and PID on the first page of any assignment turned in. In addition, student submissions must write down the names of other students who they collaborate with on the first page.

---

---

### Problem Set #1: The Fundamentals and Bias in Self-Reported Turnout

- The problem set includes warm up exercises, reviewing average treatment effects (ATE), descriptive statistics, and interpretation.
- *Due: Monday 07/07 at 11:59 PM*

### Problem Set #2: Unions, Ideology, Local Government, and Corruption

- The problem set includes treatment effects, selection into treatment, difference-in-means, regression, and covariate balance.
- *Due: Monday 07/14 at 11:59 PM*

---

### Problem Set #3: The National Supported Work Demonstration Program

- The problem set includes comparing experimental and non-experimental data, more regression, and matching.
- *Due: Monday 07/21 at 11:59 PM*

### Problem Set #4: Indiscriminate Violence and Insurgent Attacks

- The problem set includes difference-in-differences, clustered standard errors, and identification assumptions.
  - *Due: Monday 07/28 at 11:59 PM*
- 
- 

## Final Project | Research Proposal

The final project requires you to design a research proposal for a social science question of your interest using the methods and assumptions learned in this course. The goal of this assignment is to develop your ability to design a study with the intention of making policy with data while hopefully advancing your individual professional interests. You will have time in-class to develop your projects and ask questions. Problem sets will also include check points so the instructional staff can provide you feedback on your projects as you design them throughout the course period. The final submission will be 6-7 pages .pdf at 12 point font and one-half spacing. This page count does not include your cover page with abstract, reference page, figures, tables, or equations that you may incorporate in your proposal. It should include the following elements in this order:

1. **Cover Page with Abstract:** 150 word single-spaced abstract with cover page including the title of your project, your name, the abstract, the course name, and any names of people you received help from in the form of acknowledgments. Examples will be provided.
2. **Introduction and Research Question:** What is the research question you trying to answer? What causal relationship are you looking to test? What variation or pattern are you trying to explain? This section should address why we should care about your research question. What is puzzling or interesting about your question? Is it a topic of social, economic, or political importance? Have other people tried to answer this question and found conflicting or puzzling results? Will your proposal suggest something new to an established theory?
  - *“Despite the growing research exploring public opinion toward climate change, we still know little about how people respond to environmental destruction in their communities. Does exposure to deforestation change individual political preferences?”*
3. **Theoretical Answer:** What is your proposed theoretical answer to your research question? Your answer should include the causal mechanism by which change in the independent variable leads to change in the dependent variable. This section should include a literature review: what have other scholars said about your topic or question? Finding relevant literature is its own challenging task, so feel free to consider how similar research on adjacent topics (or research from different fields, like public health, law, or economics) informs your question and answer.

- 
- *“I argue that deforestation in one’s community should lead people to increasingly vote for the Democratic party because they are more likely to have a preference for pro-environment policies that protect forests. As Author 1 (2020) demonstrates, people express more pro-climate preferences after experiencing a climate disruption. In contrast, Author 2 (2020) shows the effect is conditional on partisanship...”*

4. **Ideal Experiment:** In an ideal world, we would run an experiment to answer your research question. How would you design such an experiment? What is the study population? How would you randomize treatment and what level of analysis (individual, county, something else)? Do you think all units would react to the treatment in the same way (heterogeneous treatment effects)? Discuss the ethics of your experiment.

- *“In an ideal world, I would design an experiment randomizing individual exposure to deforestation from a nationally representative sample of the United States. However, counties with more forest are likely to respond differently to deforestation than other counties. Furthermore, it is unethical to deforest random parts of the country.”*

5. **Observational Study:** Now design an observational study (non-experimental) that would enable you to evaluate your research question. For this section, you should pick one of the following four types of studies: Controlled regression, matching, difference-in-differences, or regression discontinuity. What are confounding variables or selection effects that complicate this design? What are the identification assumptions needed for your estimate to be valid? What are the limitations to your chosen method? Further details and guidance will be made available later in the course in a formalized final prompt.

- *“A linear regression would provide the relationship between the voting the Democratic candidate in 2008 and witnessing deforestation while controlling for other predictors of vote choice. However, this assumes that Democrats do not move to areas at greater risk of deforestation.”*
- *“To see whether there is a difference between voters, I can use a geographic regression discontinuity to estimate the local average treatment effect along the border between zip codes that experience deforestation and those that do not. This has the following assumptions.”*

6. **Brief Conclusion** What does your proposal hope to contribute to your area of interest? Would you want to continue working on this?

- *“As environmental shocks become common, my proposal seeks to identify whether individual partisanship shifts as a result. I hope to continue working on this as part of my senior honors thesis, research apprenticeship paper, or writing sample for graduate school applications.”*

7. **Presentation** During the last class session, students will be required to provide a 5-7 minute presentation of their final research proposal to their peers. After their presentation, their peers will have 3 minutes to provide feedback.

Some other broad questions to guide students in their writing would be to address the following: What is your research topic? What are the key terms and their definitions? What are the disagreements on this topic? Why does this topic matter? What can we learn from this project?

- *Proposal Presentation Due: Tuesday 07/29 at 11:59 PM*
- *Final Research Proposal Due: Friday 08/01 at 11:59 PM*

---

## Extra Credit

The extra credit assignment will require students to read a selected scientific study and summarize the policy evaluation described in the paper. It is completely optional, and students will face no penalty if they choose not to complete it. The value of the extra credit will be decided later in the course. As instructor, I reserve the right to remove or include extra credit opportunities.

- *Due: Tuesday 07/30 at 11:59 PM*

## GRADING POLICY

Grades in undergraduate courses are defined as follows: A, excellent; B, good; C, fair; D, poor; F, fail; I, incomplete (work of passing quality but incomplete for good cause); and IP (In Progress). The designations P (Pass) and NP (Not Pass) are used in reporting grades for some undergraduate courses. P denotes a letter grade of C- or better. A blank grade indicates no record or report of grade was received from the instructor. W is recorded on the transcript indicating the student withdrew or dropped the course sometime after the beginning of the fifth week of a quarter. Instructors have the option of assigning plus (+) and minus (-) suffixes to the grades A, B, and C. I reserve the right to curve the scale dependent on overall class scores at the end of the course.

The grade will count the assessments using the following proportions:

- **60%** of your grade will be determined by 4 weekly assignments (15% each)
- **30%** of your grade will be determined by the final project
- **10%** of your grade will be determined by attendance

Please familiarize with our institution's [Academic Regulations and Policies](#).

## POTENTIAL ONLINE SETTING

If the course must take place over Zoom or any other remote venue, lectures will be recorded for the safety of attendants and for future reviews to improve the course. Lecture recordings will not be made available to students unless instructed by the [Office for Students with Disabilities](#). However, the instructor reserves the right to make them available for review if needed.

## COURSE POLICIES

### During Class

I encourage everyone to be active and attentive during class to better not only your experience but that of your classmates and instructor. The use of electronic devices for note taking and implementing code in R will be key for this course. While laptops and other electronic devices are instrumental for this course, please refrain from using such devices for anything but activities related to the class during the class time. Eating and drinking are allowed in class but please refrain from it affecting the course such as spills, trash, or general loud noise. If the class is virtual, please keep your microphone muted during meals.

---

## Attendance Policy

Attendance will be counted by in-person completion of course polling at random times throughout each session and submission of lab assignments to Canvas. Lab assignments are graded for completion, not correctness. You are allowed 2 unexcused absences during the course. Additional absences will affect your grade unless they are excused by the instructor. Excused absences will be accommodated for personal extenuating circumstances.

## Missed/Late Assignments

Late submissions will only be graded under valid, documented, and extreme circumstances. If you know you will not be able to complete an assignment, please let me know before class session through email. If you are not able to contact me in advance, you must contact me as soon as possible. Life gets in the way, I understand the complexities of being a student. I can work with you but it is your responsibility to contact me if you need additional support.

If an extended deadline is not authorized, a missed or late assignment will be given a zero. If an extended deadline is authorized but you fail to comply, I will treat it as a missed assignment.

Assignments, besides final exams, are eligible for regrades by emailing the instructor a one-page justification. Regraded assignments could have their score increase, remain the same, or decrease.

## Academic Integrity

Students are required to comply with [UC San Diego Academic Integrity Policy](#). There is absolutely no tolerance for cheating in this course. Please do not cheat. The consequences are not worth it, ask for help when needed. Students agree that by taking this course any required assignments involving text can be subject to submission for textual similarity review to *Turnitin.com* for the detection of plagiarism. Please also abide by [UCSD's Principles of Community](#) and the [Student Code of Conduct](#).

## Student Accessibility

Students requesting accommodations for this course due to a disability must provide a current Authorization for Accommodation (AFA) letter issued by the [Office for Students with Disabilities](#). You are required to discuss accommodation arrangements with instructors and OSD liaisons in the department in advance of any exams or assignments. The OSD Liaison for the Department of Political Science is Zain Sharifi; please contact him via the Virtual Advising Center as soon as possible. I will be glad to meet with you privately during office hours to discuss accommodations.

Discrimination on the basis of race, color, national origin, religion, sex, gender, gender identity, gender expression, pregnancy (including pregnancy, childbirth, and medical conditions related to pregnancy or childbirth), physical or mental disability, medical condition, genetic information, ancestry, marital status, age, sexual orientation, citizenship, or service in the uniformed services (including membership, application for membership, performance of service, application for service, or obligation for service in the uniformed services will not be tolerated. Retaliation against any person who complains about discrimination is also prohibited. If students have questions

---

about student-related nondiscrimination policies or concerns about possible discrimination or harassment, they should contact the [Office for the Prevention of Harassment & Discrimination \(OPHD\)](#) at (858) 534-8298, [ophd@ucsd.edu](mailto:ophd@ucsd.edu), or [reportbias.ucsd.edu](mailto:reportbias.ucsd.edu).

## **Religious Accommodation**

It is the policy of the university to make reasonable efforts to accommodate students having bona fide religious conflicts with scheduled examinations by providing alternative times or methods to take such examinations. If a student anticipates that a scheduled examination will occur at a time at which his or her religious beliefs prohibit participation in the examination, the student must submit to the instructor a statement describing the nature of the religious conflict and specifying the days and times of conflict.

For final exams, the statement must be submitted no later than the end of the second week of the quarter. For all other examinations, the statement must be submitted to the instructor as soon as possible after a particular examination date is scheduled. If a conflict with the student's religious beliefs does exist, I will attempt to provide an alternative, equitable examination that does not create undue hardship for the instructor or for the other students in the class. Please see the [EPC Policies on Religious Accommodation, Final Exams, Midterm Exams](#)

## **Additional Resources**

- [CARE at the Sexual Assault Resource Center](#)
- [Counseling and Psychological Services \(CAPS\)](#)
- [UCSD Basic Needs](#)
- [Writing Hub](#)
- [Supplemental Instruction](#)
- [Tutoring](#)
- [CAPS Student Health and Well-Being](#)
- [Community Centers](#)
- [Remote Access](#)

---

## Planned Schedule and Learning Goals

The schedule is subject to change. The topics covered below should be viewed as the key concepts students should be familiar with and to use a reference for assignments. For calendar-related questions, please refer to the [Summer Session Calendar](#).

———— WEEK 1 —————

### 6/30 - Session 1: The Potential Outcomes Framework

- **Topics Covered:** Course Guidelines, Causality, The Neyman-Rubin Model, Fundamental Problem of Causal Inference
- **Reading:** Syllabus, Canvas Page, Gertler Ch 1 (Pages 3-9), Gertler Ch 3 (Pages 33-40), Imai Ch 1, Imai Ch 2 Pages (40-50)

### 7/2 - Session 2: Omitted Variable Bias and Selection

- **Topics Covered:** Difference-in-Means (DM), Before-and-After (BA) estimators, Omitted Variable Bias, Different Research Designs
- **Reading:** Gertler Ch 3 (Pages 40-47), Imai Chapter 2 (Pages 59-65)

## PROBLEM SET #1 DUE MONDAY 7/7 AT 11:59 PM

———— WEEK 2 —————

### 7/7 - Session 3: Randomized Controlled Trials I

- **Topics Covered:** Experiments, Randomization, Inference, and Ethics
- **Reading:** Gertler Ch 4 (Pages 49-69), Imai Ch 2 (Pages 51-54)

### 7/9 - Session 4: Randomized Controlled Trials II

- **Topics Covered:** Spillovers, Compliance, Interaction and Heterogeneous Treatment Effects
- **Reading:** Gertler Ch 11 (Pages 171-195)

## PROBLEM SET #2 DUE MONDAY 7/14 AT 11:59 PM

———— WEEK 3 —————

### 7/14 - Session 5: Selection on Observables and Regression

- **Topics Covered:** Confounding Variables, Linear Regression, Fixed Effects
- **Reading:** Imai Ch 4 (Pages 161-176)

---

**7/16 - Session 6: Regression and Matching**

- **Topics Covered:** Pre-processing versus Analysis, Observable Versus Unobservable Variables, Matching Designs
- **Reading:** Gertler Ch 7

**PROBLEM SET #3 DUE MONDAY 7/21 AT 11:59 PM**

———— WEEK 4 —————

**7/21 - Session 7: Difference-in-Differences and Synthetic Controls**

- **Topics Covered:** Panel Data, Within-Unit Variation, Parallel Trends Assumption, Stable Composition Assumption
- **Reading:** Gertler Ch 6, Imai Ch 2 (Pages 60-63)

**7/23 - Session 8: Regression Discontinuity Designs**

- **Topics Covered:** Treatment Threshold, Running Variables, Hard versus Fuzzy Discontinuities, Local Average Treatment Effects
- **Reading:** Gertler Ch 5

**PROBLEM SET #4 DUE MONDAY 7/28 AT 11:59 PM**

**EXTRA CREDIT DUE TUESDAY 7/29 AT 11:59 PM**

———— WEEK 5 —————

**7/28 - Session 9: Instrumental Variables**

- **Topics Covered:** Exclusion Restriction, Two Stage Least Squares Regression, and Encouragement Designs
- **Reading:** Gertler Ch 4 (Pages 69 - 79)

**7/30 - Session 10: Conclusion and Student Presentations**

- **Topics Covered:** Overview of the Course, Student Presentations, and Peer-to-Peer Feedback
- **Sign Up Sheet:** [Link Here](#)

**FINAL RESEARCH PROPOSAL DUE FRIDAY 8/1 AT 11:59 PM**