EDP 381C-14: Causal Inference
Spring 2020, Tues/Thur, 11:00 am - 12:20 pm
SZB 526

Instructor: James E. Pustejovsky (pronounced "PUHS-tea-UV-ski" or "Pusto" for short)
Email: pusto@austin.utexas.edu
Phone: 512-471-0683
Office: SZB 538D
Office hours: By appointment

Course Description

This course introduces the contemporary statistical approach to addressing questions about the causal effects of programs, policies, or interventions, focusing in particular on applied data-analysis strategies and interpretation. The course begins with an introduction to the potential outcomes framework for expressing causal quantities, followed by an examination of (idealized) simple and block randomized experiments as prototypes for learning about causal effects. The remainder of the course covers theory and data-analysis strategies for drawing causal inferences from observational studies, in which treatment conditions are not randomly assigned. Analysis techniques such as regression discontinuities, matching methods, propensity-score methods, and instrumental variables are covered both in theory and in application. Further, advanced topics will be covered based on student interest.

Learning Goals

After completing this course, students should be able to:

- Translate research questions into the framework of the potential outcomes model and specify causal quantities to be estimated.
- Articulate and assess the assumptions behind different strategies for estimating treatment effects and drawing causal inferences.
- Understand the conditions under which different causal inference strategies work well.
- Conduct, interpret, and defend a causal analysis of observational data.
- Critically review published research that addresses causal questions.

Prerequisites

- A previous course in research methods, such as EDP 381C-2 – Research Design and Methods for Psychology and Education
- A previous course in regression analysis, such as EDP 380C – Correlation and Regression
- Experience with writing scripts/programming with at least one software platform for data management and analysis (e.g., R, SAS, Stata).

Readings

- **(Required)** Gerber & Green (2012). *Field Experiments: Design, Analysis, and Interpretation*. W. W. Norton.
- **(Optional)** Pearl, Glymour, & Jewell (2016). *Causal Inference in Statistics: A Primer*. United Kingdom: John Wiley & Sons Ltd.
- Additional readings posted on Canvas.

Additional Recommended Resources

- Imbens & Rubin (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.
- Hernán & Robins (2020). *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC. Available at https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
- Cunningham (2019). *Causal Inference: The Mixtape*. Available at https://www.scunning.com/mixtape.html

Computing

In-class software demonstrations will use the R environment for statistical computing. In principle, you are welcome to complete the data-analysis exercises and the course project using your choice of software. However, I am unable to provide examples, debugging help, or technical support for any software not demonstrated in class.

There are many freely available resources for learning R. Here are some:

- R: https://www.r-project.org/
- RStudio: https://www.rstudio.com/
- TryR code school: http://tryr.codeschool.com
- Data Camp: https://www.datacamp.com/courses/free-introduction-to-r
- YaRrr! The Pirate's Guide to R: http://nathanieldphillips.com/thepiratesguidetor/
- Princeton R tutorials: http://data.princeton.edu/R/
- D-Lab R training: https://github.com/dlab-berkeley/R-for-Data-Science

Evaluation

- Data analysis exercises (70%). Students will complete problem sets involving analysis of real or simulated data. The exercises will involve implementing different analytic methods and interpreting the results.
- Course project (30%). See below.

A tentative rubric for assignment of final grades is listed below. ***The instructor reserves the right to modify this rubric.*** Square brackets correspond to ≤ or ≥; rounded parentheses to < or >.

A	[90, 100]	C+	[74, 77)
A-	[87, 90)	C	[70, 74)

B+	[84, 87)	C-	[67, 70)
B	[80, 84)	D	[60, 67)
B-	[77, 80)	F	[0, 60)

Course project

There are three options for the course project:

Option 1: Find a published study that uses one or more of the techniques discussed in the course to evaluate the causal effects of a program, policy, or intervention and ***for which the raw data is available*** (either a study that does secondary data analysis of a publically available dataset, or a study where the authors have made the raw data available through an archive or repository). Replicate the main analysis of the paper. Re-analyze the data using at least one alternative method. Critically assess the findings of the study.

Option 2: Conduct an observational study using one of the causal inference methods discussed in the course. Submit a paper presenting the results of the study and covering: the research question, data, empirical strategy, results, and conclusions. The data-analysis code should be submitted as an appendix. You are free to choose any topic you like, as long as you have a clear research question that concerns the causal effect of some intervention, treatment, policy, or event on some outcome, result, or performance.

Option 3: Pick a policy topic and critically evaluate the empirical evidence about a particular program or set of programs, covering at least three relevant studies. Write a report in which you: 1) discuss and carefully define the causal questions relevant to assessment of the program(s); 2) discuss and carefully define the questions that the studies attempt to answer; 3) describe and discuss the appropriateness of the study designs; 4) describe the results; and 5) assess the state of knowledge about your policy topic based on your chosen studies and identify knowledge gaps.

Attendance

Students are responsible for all of the material presented during class meetings. If a student must miss a class, it is their responsibility to obtain from classmates and thoroughly review notes or summaries of the material that they missed.

Academic integrity and plagiarism

Following the University's honor code, students are expected to maintain absolute integrity and a high standard of individual honor in scholastic work. Assignments and projects must be completed with the utmost honesty, which includes acknowledging the contributions of other sources to your scholastic efforts; avoiding plagiarism; and completing assignments independently unless expressly authorized otherwise. ***Homework assignments or projects containing any plagiarized material will not be accepted.***

Carrying of Handguns

Students in this class should be aware of the following university policies:

- Individuals who hold a license to carry are eligible to carry a concealed handgun on campus, including in most outdoor areas, buildings and spaces that are accessible to the public, and in classrooms.
- It is the responsibility of concealed-carry license holders to carry their handguns on or about their person at all times while on campus. Open carry is NOT permitted, meaning that a license holder may not carry a partially or wholly visible handgun on campus premises or on any university driveway, street, sidewalk or walkway, parking lot, parking garage, or other parking area.

ADA accommodations

The University of Texas at Austin provides upon request appropriate accommodations for qualified students with disabilities. For more information, please contact the Office of the Dean of Students at 471-6259, 471-4671 TTY.

Religious Holidays

By UT Austin policy, students must notify the instructor of a pending absence due to religious observance at least fourteen days in advance. If the student must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, the student will be given an opportunity to complete the missed work within a reasonable time after the absence, with no penalty.

Emergency Evacuation Policy

Occupants of buildings on the UT Austin campus are required to evacuate and assemble outside when a fire alarm is activated or an announcement is made. Please be aware of the following policies regarding evacuation:

- Familiarize yourself with all exit doors of the classroom and the building. Remember that the nearest exit door may not be the one you used when you entered the building.
- If you require assistance to evacuate, inform the instructor in writing during the first week of class.
- In the event of an evacuation, follow the instructions of the instructor.

Do not re-enter a building unless you're given instructions by the Austin Fire Department, the UT Austin Police Department, or the Fire Prevention Services office.

Course Outline

This outline is tentative and will almost certainly evolve over the course of the semester.

1/21 – Course introduction

- Holland (1986)
- Rubin (1986)
- (Optional) Hernan & Robins (2020), chapter 2

1/23 – Probability crash course

- Course notes on probability (sections 1 & 2)

1/28 – Probability (continued)

- Course notes on probability (section 3)

1/30 – Simple randomized experiments

- Gerber & Green (2012), chapters 1-2, 3.1-3.5
- (Optional) Hernan & Robins (2020), chapter 1

2/4 – Block-randomized experiments

- Gerber & Green (2012), chapters 3.6, 4

2/6 – Balance assessment & covariate adjustment

- Gerber & Green (2012), chapter 4

2/11 - Attrition

- Gerber & Green (2012), chapter 7

2/13 – Block-randomized experiments (revisited)

2/18 – Instrumental variables

- Gerber & Green (2012), chapter 5
- (Optional) Kim et al. (2011).

2/20 – Instrumental variables (continued)

- Gerber & Green (2012), chapter 6
- (Optional) Angrist, Imbens, & Rubin (1996)

2/25 – Regression discontinuities

- Bloom (2012)
- Skovron & Titiunik (2015)

2/27 – More on regression discontinuities

3/3 – Observational studies and confounding, directed acyclic graphs (DAGs)

- Pearl, Glymour, & Jewell (2016), chapters 1 & 2.

3/5 – Identification with DAGs

- Pearl, Glymour, & Jewell (2016), chapter 3.

- (Optional) Elwert (2013).

3/10 – Regression analysis
- Schafer & Kang (2008)

3/12 – Regression analysis (continued)
- Ho, King, Imai, & Stuart (2007)

3/17 – No class (Spring break)

3/19 – No class (Spring break)

3/24 – Confounding on a single covariate
- (Optional) Cochran & Rubin (1973)

3/26 – Multivariate matching using propensity scores
- Schafer & Kang (2008)
- Rosenbaum & Rubin (1983)

3/31 – Propensity score analysis: stratification and matching
- Stuart (2010)
- Ho, King, Imai, & Stuart (2007)

4/2 – Propensity score analysis: weighting
- Schafer & Kang (2008)
- (Optional) Hirano & Imbens (2001)

4/7 – Balance assessment
- Greifer (2017).

4/9 – Machine Learning
- McCaffrey, Ridgeway, & Morral (2004)

4/14 – Machine Learning (continued)
- Hill (2011).

4/16 – Difference-in-differences/fixed effects models
- Angrist & Pischke (2009).

4/21 – More difference-in-differences

4/23 – Causal mediation
- Vanderweele & Vansteelandt (2009).

4/28 – Causal mediation (continued)

4/30 – TBD

5/5 – TBD

5/7 – TBD