Causal Inference, 2 Days Workshop

Dr. Florian Foos
London School of Economics & Political Science f.foos@lse.ac.uk

Information: "Correlation is not causation". You must have heard this warning many times. But, what then is causation, and how can we identify the causal effect of political messages, policies or programmes? This module introduces students to the ideal of the randomized experiment, and other common methods of causal inference using quasi-experimental data.

Aims: The module provides an introduction to the design-based approach to causal inference. Instead of using fancy modelling approaches to correct post-hoc for potential biases, the module encourages students to think about challenges to causal inference at the design stage of a study. Published work will be evaluated based on how it addresses three key assumptions underlying causal inference: independence, excludability, and non-interference. The module will cover identification strategies for a variety of causal research questions in Political Science. After introducing participants to the design-based approach to causal inference in general, the module will cover the Regression Discontinuity Design and the Difference-in-Differences Design, two methods that allow design-based inference with quasi-experimental data, in detail. The module is taught as a combination of lectures and applied computer labs in Stata (R code will be provided).

Prerequisites: The only pre-requisite is any course covering linear regression. There is relatively little assumed knowledge, and the aim is to build the statistical foundations from the ground up. If you have conducted a hypothesis test of any kind, you probably have the requisite skills.

Learning Outcomes: Participants will understand the potential outcomes framework, and the key assumptions underlying causal inference, and will be able to choose appropriate methods for a variety of research questions posing different identification challenges. Moreover, they will gain the practical skills of applying these insights and the statistical knowledge to their own research ideas.

Course Dropbox: Readings and code files are available from the course dropbox.

Required textbook
Cunningham, Scott. *Causal Inference: The Mixtape. V1.7*, 2019: available via course dropbox.

Recommended textbooks
Angrist, Joshua and Joern-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricists Companion.*, Princeton: Princeton University Press, 2009.

Bueno de Mesquita, Eitan, and Anthony Fowler. *Thinking Clearly in a Data-Driven Age*, 2019: available via course dropbox.

Dunning, Thad. *Natural Experiments in the Social Sciences. A Design-Based Approach*, Cambridge: Cambridge University Press, 2012.

Gerber, Alan and Donald P. Green. *Field Experiments: Design, Analysis, and Interpretation*, New York: W.W. Norton, 2012.

Imbens, Guido W. and Donald B. Rubin. *Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction.* Princeton: Princeton University Press, 2015.

Software: Participants will have a choice between using STATA (demonstration in class) and R (code will be provided).

Workshop Outline

1. Introduction to causal inference
2. Potential Outcomes Framework and ATE
3. Regression Discontinuity Design (RDD): Theory
4. Regression Discontinuity Design (RDD): Application
5. Difference-in-Differences Design (DiD): Theory
6. Difference-in-Differences Design (DiD): Application
7. Advice sessions on individual research designs

Introduction

Reading

Gelman, Forward Causal Inference

Gerber and Green: Chapters 1 and 2.

Content

- What is causal inference?
- Unobserved heterogeneity
- Experiments and quasi-experiments
- Random and quasi-random assignment
- Potential outcomes notation
- The three core assumptions

Fundamental problem of causal inference and the ATE

Reading

Cunningham: p.81f "Potential outcomes causal model"

Gerber and Green: Chapter 2

Holland, Fundamental Problem of Causal Inference

Content

- Fundamental problem of causal inference
- Potential and realised outcomes
- The switching equation
- Difference-in-means estimator of the ATE
- Unbiasedness of ATE estimator
- Sampling distribution of the ATE

Regression Discontinuity Design (RDD)

Reading - basics

Cunningham: p.153f "Regression discontinuity'

Bueno de Mesquita and Fowler, p. 338f

Angrist and Pischke, chapter 6

Dunning, chapters 1 and 3

Campbell, Donald T. 1969. Reforms as experiments. *American Psychologist* 24(4): 409-429.

Reading - advanced

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of econometrics* 142(2): 615-635.

Caughey, Devin, and Jasjeet S. Sekhon. 2011. "Elections and the regression discontinuity design: Lessons from close US house races, 19422008." Political Analysis 19(4): 385-408.

Keele, Luke, and Rocio Titiunik. 2016. "Natural experiments based on geography." *Political Science Research and Methods* 4(1): 65-95.

Content

- What is a quasi-experiment
- RDD assumptions
- The running and assignment variables
- Polynomial estimation
- Different bandwidth estimators
- Robustness checks: Balance tests, density tests, placebo outcomes
- Fuzzy RDD
- Rdrobust and rdpower packages

Difference-in-differences design (DiD)

Reading

Cunningham: p.263f "Difference-in-differences"

Bueno de Mesquita and Fowler, p. 355f

Angrist and Pischke, chapter 5

Imai, chapter 2

Content

- Basics
- Parallel trends assumption
- Confounding
- Extension to multiple groups and treatment periods
- Examples: Hillsborough and Euroscepticism (Foos and Bischof 2019); Craigslist and Violence against Women (Cunningham 2019)
- diff package