Causal Inference from Observational Data
Instructor: Felix Elwert, Ph.D.

Short Description:

This course offers an applied introduction to causal inference for social scientists. The content is structured around four key topics: Modern causal notation (potential outcomes and directed acyclic graphs [DAGs]), propensity score matching, instrumental variables, and inference for time-varying treatments. Subsidiary topics include the art of formulating causal questions, the mechanics of balance testing, sensitivity analysis, effect heterogeneity, social networks, and causal mediation analysis. Throughout, the course will focus on building strong transferable intuition. We will prioritize a thorough substantively grounded understanding of assumptions over mathematical proofs and derivations. The goal is to empower applied social scientists to apply new causal tools with confidence.

Course Outline

Segment 1: Causality and Potential Outcomes
The modern literature on causality begins with Rubin's (1974) definition of causal effects as differences between potential outcomes. The potential outcomes framework is now the dominant framework of causal inference. Its major contribution was not the introduction of new techniques (of which there are few) but the clarification of conventional practices. This segment introduces potential outcomes and teaches students to ask properly formulated causal questions.

The counterfactual framework

- Causal effects as differences between potential outcomes
- The Fundamental problem of causal inference
- Randomized experiments
- Observational studies

Complications

- Manipulability, SUTVA, and effect heterogeneity

Asking causal questions

- Group exercises.

Segment 2: Directed acyclic graphs (DAGs)
Directed acyclic graphs are an equivalent and often more accessible alternative to potential outcomes notation. This segment introduces the main uses of DAGs for notating and working with causal models.

DAGs: The Basics

- Notation: nodes, edges, missing edges
- Navigating between causation and association
- d-Separation:
  - Causation
  - Confounding
  - Endogenous selection

Testable Implications of a causal model

- Deriving *all* testable assumptions of a causal model.

Identification

- Identification: when does association imply causation?
- Identification in regression and matching: Pearl's backdoor criterion
- Examples

Endogenous selection bias

- A complete enumeration of structural biases in causal inference
- Selection bias as "conditioning on a collider variable"
- Lots of real examples across the social sciences

Segment 3: Exact and propensity score matching

Whereas segments 1 and 2 were dedicated to a deep understanding of conceptual underpinnings with lots of real examples (thinking!), Segments 3-5 emphasize the how, why, and when of actual estimation (doing!). Matching as presently perhaps the most popular strategy for estimating causal effects. Students will learn the logic and intuition of exact matching, propensity score matching, and their relationship to regression.

Exact matching

- Experimental analogy:
- Examples
- Curse of dimensionality

Propensity score matching

- A solution to the curse of dimensionality
- Matching options: blocking, caliper, kernel, etc
- Estimation options: parametric, nonparametric, regression, weighting, doubly robust estimation.
- Examples

Balance testing (1h)

Balance testing is the oft-neglected central part of any matching exercise. Matching is simply a means to the end of creating balance.

- Balance testing fallacy
- Graphical and numerical tests

Sensitivity analysis and bounds
- Formal sensitivity analysis
- Manski Bounds

Effect heterogeneity
- Simple effect heterogeneity
  - How propensity score matching is immune
  - How regression recovers only variance-weighted average effects
- Pervasive effect heterogeneity
  - Heterogeneity as confounding bias
- Examples

Segment 4: Instrumental variables
Instrumental variables (IV) are a solution to the omitted variables problem. This segment focuses on the complementary presentations of IV (algebra, DAGs, and potential outcomes) in learn and nonlinear models. We emphasize a substantive understanding an testing of IV assumptions and promising uses.

Basic IV
- Assumptions
- Non-parametric test of the null hypothesis
- Parametric estimation in linear models

How to find IVs
- Graphical IV criterion

Understanding assumptions and detecting violations
- Weak instrument bias
- Exclusion violations
- Formal and informal tests

IV for social network analysis

Segment 5: Mediation and time-varying treatments
Some of the most exciting work on causal inference over the past decade deals with time-varying treatments. This includes causal mediation analysis. This segment conveys dependable intuitions for the underlying problems and the shortcomings of existing methods, and it gives pointers for new solutions.

Identification problems
- Dealing with post-treatment variables

Mediation analysis: concepts and identification

- Controlled direct effects
- Natural direct and indirect effects
- Graphical identification criteria for mediation analysis
- Why randomized experiments cannot save the day

Time-varying treatments

- Static vs dynamic estimands
- Identification: sequential ignorability with graphs
- Inverse-probability of treatment weighting
- Marginal structural models

Concluding discussion

Requirements
Participants should have a solid applied background in multiple regression. The course does not require calculus or matrix algebra, though neither will hurt. We will execute empirical examples in Stata.

Instructor
Felix Elwert (Ph.D., Harvard 2007) is a Vilas Associate Professor of Sociology at the University of Wisconsin-Madison. He is the winner of the first Causality in Statistics Education Award, given by the American Statistical Association in 2013. His research focuses on methods of causal inference with applications in social stratification, social demography, and network analysis. His work has appeared in the American Journal of Sociology, the American Sociological Review, Demography, the American Journal of Public Health, and Biometrics. He regularly teaches courses on causal inference in the United States and Europe.