

IDEM 185

Advanced Methods for Causal Inference with Observational Data

Course coordinator: Peter Eibich

Start date: 19 July 2021

End date: 23 July 2021

Location: Online Course. Link tba.

Instructors

- Peter Eibich
- Angelo Lorenti

Course description

Many research questions in demography, social sciences, economics, or epidemiology deal implicitly or explicitly with causal effects or causal mechanisms. For example:

- How does education affect the timing of fertility?
- How does health affect labor market participation?
- What is the impact of overweight and obesity on healthcare utilization?

Unfortunately, conducting randomized experiments to study these questions is often infeasible, and we therefore rely on research designs and statistical and econometric methods to identify causal effects under certain assumptions.

In this course, we will review selected research designs and estimation methods that can be used to obtain causal estimates. Over the course of five days, we will review the theory behind these estimation methods and implement them in a series of practical exercises using real-world data. We will also discuss the research designs and relevant assumptions employed in published papers to identify causal effects.

Organization

The course will be offered online. On Monday through Thursday, there will be one lecture and one practical exercise / computer lab session. All lectures will be pre-recorded and made available on Thursday, July 15, four days before the course. Students are expected to watch the lecture carefully before the practical exercise / computer lab session, which will be held live every day from 15:30-17:30 CEST (Central European Summer Time). During the daily

practical exercise /computer lab sessions, students will gain hands-on experience. In each session, students and instructors will review code together, perform hands-on exercises, and evaluate these exercises together. This will allow students to develop intuition in using the concepts and deepen their understanding of the concepts.

On Friday from 15:30-20:30 CEST, students are expected to make a presentation (see below, "Examination").

In general students should expect to spend about 6-8 hours per day on the course (lectures, practical exercise sessions, readings, presentations).

Detailed schedule

The course will cover the following topics in this order.

Monday

1. Advanced Instrumental Variables estimation
 - a. Testing IV assumptions
 - b. Linear and nonlinear IV estimation
 - c. Characterizing the complier population

Tuesday

2. Regression Discontinuity Designs
 - a. Sharp and Fuzzy RDDs
 - b. Visual analysis
 - c. Bandwidth choice and estimation

Wednesday

3. General Introduction to Propensity Scores (PS) and Applications
 - a. Methods for PS Estimation
 - b. Implementation
 - c. Covariate Balance Evaluation

Thursday

- d. Treatment Effect Estimation
- e. Sensitivity Analysis

Course prerequisites

A working knowledge of STATA or R is recommended for the practical exercises. Students working with other statistical software packages can participate in the course, but support for the practical exercises will only be provided for STATA and R.

The course content will build on the topics covered in the [PHDS Population Health course](#) and assumes basic knowledge of linear and non-linear regression models as well as instrumental variables estimation.

Students who have not attended [PHDS Population Health course](#) should make sure that they are familiar with the following material:

- Angrist, J., Pischke, J.-S., 2009. Mostly harmless econometrics. Princeton University Press, Princeton and Oxford. (Chapters 3 & 4)
- Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.
- Students who do not have access to the Imbens & Rubin (2015) book should study this [freely available online book](#): Hernán M.A., Robins J.M. (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC. (Chapters 2, 12 & 15)

Examination

Towards the end of the course students will be asked to give a brief presentation on a published research paper of their choice. The presentation should consist of a brief overview of the paper as well as reflections on the causal research design (or lack thereof). Further details on the format of the presentation will be shared closer to the date.

General readings

Linear Regression Basics and Instrumental Variables:

Angrist, J., Pischke, J.-S., 2009. Mostly harmless econometrics. Princeton University Press, Princeton and Oxford. (Chapters 3 & 4)

A broad overview of Regression Discontinuity Designs:

Lee, D.S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. Journal of Economic Literature 48, 281–355. <https://doi.org/10.1257/jel.48.2.281>

General Introduction to Propensity Score Methods:

Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.

Admission

This course is restricted to IMPRS-PHDS students and MPIDR scientific staff.

Questions about the course should be directed to phds@demogr.mpg.de