THE STATISTICS OF CAUSAL INFERENCE

Organised by Elias Dinas, Andrea Ichino & Miriam Golden
20-lectures
September 2020 — January 2021
Thursdays and Tuesdays 15-17
Opening lecture: September 17, 2020
Online via Zoom
40 credits

Deadline for registration for SPS: 23 September 2020

Course Outline

The objective of this course is to learn how statistical methods can help us to draw causal claims about phenomena of interest. Students will be introduced into the two authoritative frameworks of causal inference in social sciences, i.e. the Potential Outcomes framework and the Directed Acyclic Graphs framework. By the end of the course, students will be in position to: a) critically read and evaluate statements about causal relationships based on some analysis of data; b) apply a variety of design-based easy-to-implement methods that will help them draw causal inferences in their own research; and think about data under the logic of causal inference. This is a joint course between Econ and SPS and assumes basic knowledge of linear regression.

For the SPS department, the course will give 40 credits to students attending all lessons and passing the final exam. Students are allowed to attend also half of the lectures (at their choice) in which case the course will give 20 credits.

For the ECO department, the course will give 2 credits to students attending all lessons and passing the final exam. Students are allowed to attend also half of the lectures (at their choice) in which case the course will give 1 credit.

1. **Introduction to Causality (1 lectures)**
   ANDREA, ELIAS, MIRIAM
   Thursday September 17, 15.00-17.00

   - The problem and examples

- The potential outcomes framework
- Why the naive estimator is not a solution
- Sample selection bias

2. **Randomized experiments (1 lecture)**
   ELIAS
   Tuesday September 22, 15.00-17.00

   - The best solution with some cons: theory, external validity
   - Randomization inference
   - Randomized control trials
   - Examples

3. **Standard regression analysis and causality (1 lecture)**
   ANDREA
   Thursday September 24, 15.00-17.00

   - The Population Regression function
   - Conditions under which it has a causal interpretation
   - RCT and regression
   - Alternatives when Randomization is not an option: the rest of the course

4. **Difference in difference (3 lectures)**
   ELIAS
   Tuesday September 29, 15.00-17.00
   Thursday October 1, 15.00-17.00
   Tuesday October 6, 15.00-17.00

   - The identification strategy
   - What must be true for the strategy to work: pros and cons
   - Implementation
   - Examples

5. **Solutions based on control functions (2 lectures)**
   ANDREA
   Thursday October 8, 15.00-17.00
   Tuesday October 13, 15.00-17.00

   - Heckman two step procedure
   - Generalized control function approach
   - What must be true for the strategy to work: pros and cons
   - Examples

6. **Instrumental variables (2 lectures)**
   ANDREA
   Thursday October 15, 15.00-17.00
   Tuesday October 20, 15.00-17.00

- The traditional intepretation of IV
- The LATE interpretation of IV
- What must be true for the strategy to work: pros and cons
- Examples

7. **Regression discontinuity designs (3 lectures) ELIAS (Andrea on multiple threshold and some examples from his work)**
   ELIAS, ANDREA
   Thursday October 22, 15.00-17.00
   Thursday October 29, 15.00-17.00
   Tuesday November 10, 15.00-17.00

   - The identification strategy
   - Sharp RDD
   - Fuzzy RDD and the local LATE interpretation of RDD
   - What must be true for the strategy to work: pros and cons
   - The actual implementation and its problems
     - Testing identification conditions
     - Multiple thresholds
   - Examples

8. **Experiments (1 lectures)**
   MIRIAM
   Thursday November 12, 15.00-17.00

   - An encompassing logic of design-based causal inference with observational data
   - Examples and Applications
   - What can go wrong

9. **Matching methods (1 lecture)**
   ANDREA
   Tuesday November 17, 15.00-17.00

   - The identification strategy
   - What must be true for the strategy to work: pros and cons
   - Exact matching
   - Propensity score matching
   - Examples

10. **Synthetic control methods (1 lecture)**
    ELIAS
    Thursday November 19, 15.00-17.00

    - The identification strategy
    - What must be true for the strategy to work: pros and cons
    - examples

11. **Mediation Analysis (2 lecture)**
    ELIAS
    Thursday November 26, 15.00-17.00

    - Introduction to Directed Acyclic Graphs
    - The back-door criterion
    - The front-door criterion
    - What must be true for the strategy to work: pros and cons
    - examples

12. **Uncertainty (1 lectures)**
    ELIAS
    Thursday December 3, 15.00-17.00
    Thursday December 10, 15.00-17.00

    - Parametric Inference
    - Non-parametric inference: resampling with replacement
    - Randomization inference
    - Clustering

13. **Wrap Up (1 lecture)**
    ANDREAS, ELIAS, MIRIAM
    Thursday December 17, 15.00-17.00

    - Summary, applications and extensions

Exercise classes
No exercise class

Teaching material (available on Brightspace or in the library)

- Lecture notes by the instructor
- Josh Angrist, Steve Pischke, "Mostly harmless econometrics: an empiricist's companion", Princeton University Press,
- Josh Angrist, Steve Pischke, "Mastering metrics: the path from cause to effect", Princeton University Press,
- Articles in journals

Final exam and Grading
There will be final class room exam (referee report) and a take home exam (simulation exercise).