Causal Inference and Research Design

Instructor: Scott Cunningham

Description

This online course will provide a foundation and review of a statistical framework within econometrics and other branches of the social sciences called "causal inference." Causal inference is a practice which attempts to determine whether given two events, one event caused the other. It is commonly used in program evaluation as well as research aimed to evaluate the empirical content of certain scientific theories such as estimates of the price elasticity of demand and returns to schooling. This class is meant to be a primer and so will cover the potential outcomes model, directed acyclic graphs, regression discontinuity, instrumental variables, difference in differences, synthetic control, and matching. It will be accompanied by efforts to introduce students to basic practices in programming as well as good research practices more generally.

Aims of the class:
1. To help students become more familiar with the field of causal inference
2. To empower students to apply research designs more competently to their own research
3. To direct students towards better programming practices so that they are better able to perform quantitative forms of research

Expectations:

Participants must be self-motivated and commit to use the resources provided to them to grow and perfect their own understanding of the material as well as solve programming challenges on their own. If attendees devote themselves to mastering this material to the best of their ability such that they maximize the opportunity provided by this program, it is the hope that the class will prove to be valuable in their own education, growth, and ultimately scientific potential.

Virtual Platform:

Links to the online platform will be provided a couple of days prior to the course start date.

Online Schedule:

For the online courses, each day will include lecture for 75 minutes followed by a 15-minute break. This will continue until a one-hour lunch. Then it will continue until the end of the day. Each day is expected to take eight hours.

Day one
- Introductions
- Hidden curriculum material

Causal Inference and Research Design

- o Mental health, success, and values
- o Programming, organization, workflow and research and professional development
- Potential outcomes
  - o Randomization and selection bias
  - o Fisher's sharp null (i.e., randomization inference)
- Introduction to directed acyclic graphs
- Sharp regression discontinuity

Day two:
- Instrumental variables
  - o Some two step estimators (Wald, two sample IV, two stage least squares)
  - o Minimum distance estimators (limited information maximum likelihood)
- Just identified versus multiple instruments
  - o weak and strong instruments, bias, and consistency of 2SLS
  - o Multiple instrument estimators: LIML, jackknife, unbiased jackknife, and double selection methods
- Heterogenous treatment effects and the local average treatment effect
- Four basic IV designs
- Introduction to difference in differences

Day Three
- Conclude Difference in differences
- Synthetic control
  - o Abadie, Diamond and Hainmueller canonical model
  - o Cherry picking synthetic controls
  - o Doudchenko and Imbens
  - o Augmented synth (Rothstein et al 2021)
- Matching (if possible)
  - o Backdoor criterion
  - o Approximate matching
  - o Exact matching
  - o LaLonde 1986, Dehejia and Wahba 2002