Applied Microeconometrics

Lecture 4a: Empirical strategies (ii) (‘refresher lecture’)

Manuel Bagues

August 2022

Lecture Slides

Overview

Previous lecture: Empirical strategies (‘refresher lecture’)

- Introduction
  - Structural and treatment effect approaches
  - Guidance of theory
  - ‘Identification strategy’ vs. ‘estimation method’
- Identification strategies
  - RCTs
  - Selection based on observables
  - Instrumental variables

Today

- Identification strategies
  - IV
  - Canonic DID
- Advances in DID
- Class presentation

Why IV works?

- Intuitive idea behind IV is as follows:
  1. **Relevance**: You found a variable (the instrument) that affects who is assigned to the treatment
  2. **Exogeneity**: The instrument is as good as randomly assigned
  3. **Exclusion restriction**: And you know that your instrument only affect the outcome through its impact on the treatment (cannot be correlated with any other factor that affects the outcome, or affect directly the outcome).
- Note that an IV strategy is equivalent to an RCT where there is no full compliance

Can we test validity of IV?

- Can we test the assumptions required for the validity of a IV:
  1. Is the instrument correlated with the treatment? [Cov$(Z_i, S_i) \neq 0$]
     - YES: Significance of first stage, F-statistics
  2. Exogeneity [$\eta_i \not\rightarrow Z_i$]
     - SOMETIMES YES: Is the instrument as good as random?
  3. Exclusion restriction? [$\eta_i \not\leftarrow Z_i$]
     - Is it plausible that the instrument only affects the outcome variable through its impact on the treatment?
     - We can try to show that it is not satisfied looking at the impact of the instrument on the outcome variable among always-takers or never-takers
     - But we will never be able to proof that there is no unobserved channel through which the instrument affects the outcome

Three ways to calculate the IV estimator

1. Standard formula: $\hat{\beta}_{IV} = [Z'X]^{-1}Z'Y$
2. $\hat{\beta}_{IV} = \frac{\hat{\beta}_{\text{reduced form}}}{\hat{\beta}_{\text{1st stage}}}$
3. Two-stage least squares (2SLS)
    - 1st stage: regress the treatment on the instrument (+ covariates)
    - 2nd stage: regress the outcome variable on the predicted treatment (+ covariates)

Local average treatment effects

- The impact of the treatment might heterogenous
- Therefore, it is important to understand the type of treatment effect that we are identifying
- Let $\mathbf{D_{i1}} \in \{0, 1\}$ indicate whether an individual **affected** by the instrument would receive or not the treatment. Similarly, let $\mathbf{D_{i0}} \in \{0, 1\}$ indicate whether an individual **not affected** by the instrument would receive or not the treatment
- We can consider four instrument-dependent subgroups, defined by the manner in which members of the population react to the instrument:
  - Compliers: $D_{i1} = 1, D_{i0} = 0$
  - Always-takers: $D_{i1} = 1, D_{i0} = 1$
  - Never-takers: $D_{i1} = 0, D_{i0} = 0$
  - Defiers: $D_{i1} = 0, D_{i0} = 1$

- Typically we assume that there are no defiers (*monotonicity assumption*)
  - Example: impact of promotions using connections in committees as an instrument
- Under this assumption, with an instrumental variable strategy we learn about the impact of the treatment for the group of compliers.
- IV strategy identifies the Local average treatment effect (LATE), in the sense that we learn about the impact of the treatment for a very particular group of individuals.

Example: The Effect of Peer Salaries on Job Satisfaction

- Let us consider a simple set up where there is a binary instrument and a binary treatment:
  - Card et al. 2012 "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction"
- Email to random sample of employees of University of California with link to a webpage with information on salaries.
- treatment: information; outcome: job satisfaction; instrument: email
- Impact of the instrument on the treatment:
  - People that receive the information email: 50% check the webpage
  - People that do not receive the information email: 20% check the webpage

Example: The Effect of Peer Salaries on Job Satisfaction

- Two questions
  1. Explain verbally who are the compliers, always-takers, never-takers and defiers.
  2. Let us assume that there are no defiers. What is the share of always-takers, never-takers and compliers?

Weak instruments

- Andrews, Stock and Sun (2019), 'Weak Instruments in Instrumental Variables Regression: Theory and Practice'
- Abstract: When instruments are weakly correlated with endogenous regressors, conventional methods for instrumental variables (IV) estimation and inference become unreliable. A large literature in econometrics has developed procedures for detecting weak instruments and constructing robust confidence sets, but many of the results in this literature are limited to settings with independent and homoskedastic data, while data encountered in practice frequently violate these assumptions. We review the literature on weak instruments in linear IV regression with an emphasis on results for nonhomoskedastic (heteroskedastic, serially correlated, or clustered) data. To assess the practical importance of weak instruments, we also report tabulations and simulations based on a survey of papers published in the American Economic Review from 2014 to 2018 that use IV. These results suggest that weak instruments remain an important issue for empirical practice, and that there are simple steps that researchers can take to better handle

Roadmap so far:

- With every research question it is not possible to run a randomized controlled trial.
- Maybe we can look for an instrumental variable, but good instruments are difficult to find...
- We may also try to learn about the impact of a treatment using an empirical strategy based on observables:
  - We can compare individuals exposed to the treatment with other individuals that look alike in terms of observables.
  - Unfortunately, this evidence is likely to be subject to selection biases and often it is difficult to interpret.
- What else can we do $\rightarrow$ Difference-in-differences
  - We look for a control group such that its evolution provides a good counterfactual for how the treatment group would have evolved in the absence of the treatment
  - Note: the control is assumed to evolve similar, but does not need to be similar in levels to the treatment group

Classic example: Card and Krueger (1994)

Effect of Minimum wages on employment

- On April 1, 1992, New Jersey raised the state minimum from \$4.25 to \$5.05, whereas in the bordering state of Pennsylvania the minimum wage stayed at \$4.25 throughout this period.
- Card and Krueger (1994) evaluated the effect of this change on the employment of low wage workers.
- They conducted a survey to some 400 fast food restaurants from the two states just before the NJ reform, and a second survey to the same outlets 7-8 months after.

Differences-in-differences (dif-in-dif)

- The above example captures the main intuition behind the differences-in-differences analysis.
- We use the evolution of the outcome variable in the control group to construct a counterfactual of what would have happened in the treatment group in the absence of the treatment.
- Parallel trends assumption: The fundamental identifying assumption is that, in the absence of the treatment, both groups would have followed parallel trends
- Note that this empirical strategy allows for the existence of time-invariant differences between the two groups, but it assumes that there are no time-variant relevant differences.

Main threats to validity of dif-in-dif estimates

1. If the groups are different in levels, maybe they evolve differently?
2. Why did the treatment group adopt the policy, and not the control group?
3. Policies are usually implemented in bundles (the timing of the treatment may not be by chance) $\rightarrow$ the outcome variable may be affected by these other policies
4. The treatment should not affect the control group (SUTVA)
5. No anticipation effect
6. The composition of the treatment and control groups should not change as a result of treatment

Usual checks

1. The two groups evolved similarly in the past (although note that this is neither a sufficient nor a necessary condition for the validity of the empirical strategy!)
   - Show it graphically
   - Report in a table the estimates from an event study estimation: interact ‘treatment dummy’ with lags and forwards of the ‘ time dummies’ (e.g. year dummies)
2. Make sure that the timing of the adoption of the policy was as good as random (e.g. not driven by expectations of economic shocks)
3. No other policies were adopted at the same time
4. Verify that there is no reason to belief that the control group might be affected by the treatment

Final comments on diff-and-diff:

- Identification relies on a claim that is very often more of an act of faith than a assumption clearly grounded on theory
- Show that in the past trends were parallel (it does not guarantee that this will be satisfied in the future, but at least it is supportive)
- Discuss explicitly why it is a good assumption to believe that the timing of the treatment/policy was as good as random
- Discuss explicitly the existence of alternative policies that might contemporaneously affect the treatment or the control group
- Discuss the possibility that the control group is affected by the treatment.

Combining IV with other identification strategies

- Several ways to find a valid control group:
  - RCT
    - Eg: olive oil
  - Identification based on observables (+controls)
    - Eg: fast food restaurants
  - DID (+unit and time fixed effects)
    - Eg: minimum length of education
  - RDD (above and below threshold)
    - Eg: quotas in electoral lists

- Sometimes not everybody assigned to treatment receives treatment
- To estimate the impact of the treatment, we add an IV
    - RCT+IV
    - Identification based on observables + IV
    - DID+IV
    - RDD+IV
- In addition to exegoneity, now we also need the exclusion restriction to hold
- LATE