Difference-in-Differences

Mixtape Session

MIXTAPE SESSIONS

Roadmap

Workshop outline

- Introduction to DiD basics
  - → Potential outcomes review
  - → DiD formula
  - → Covariates
- Differential timing
  - → Heterogeneity
  - → TWFE bias in estimation of overall and dynamic ATT

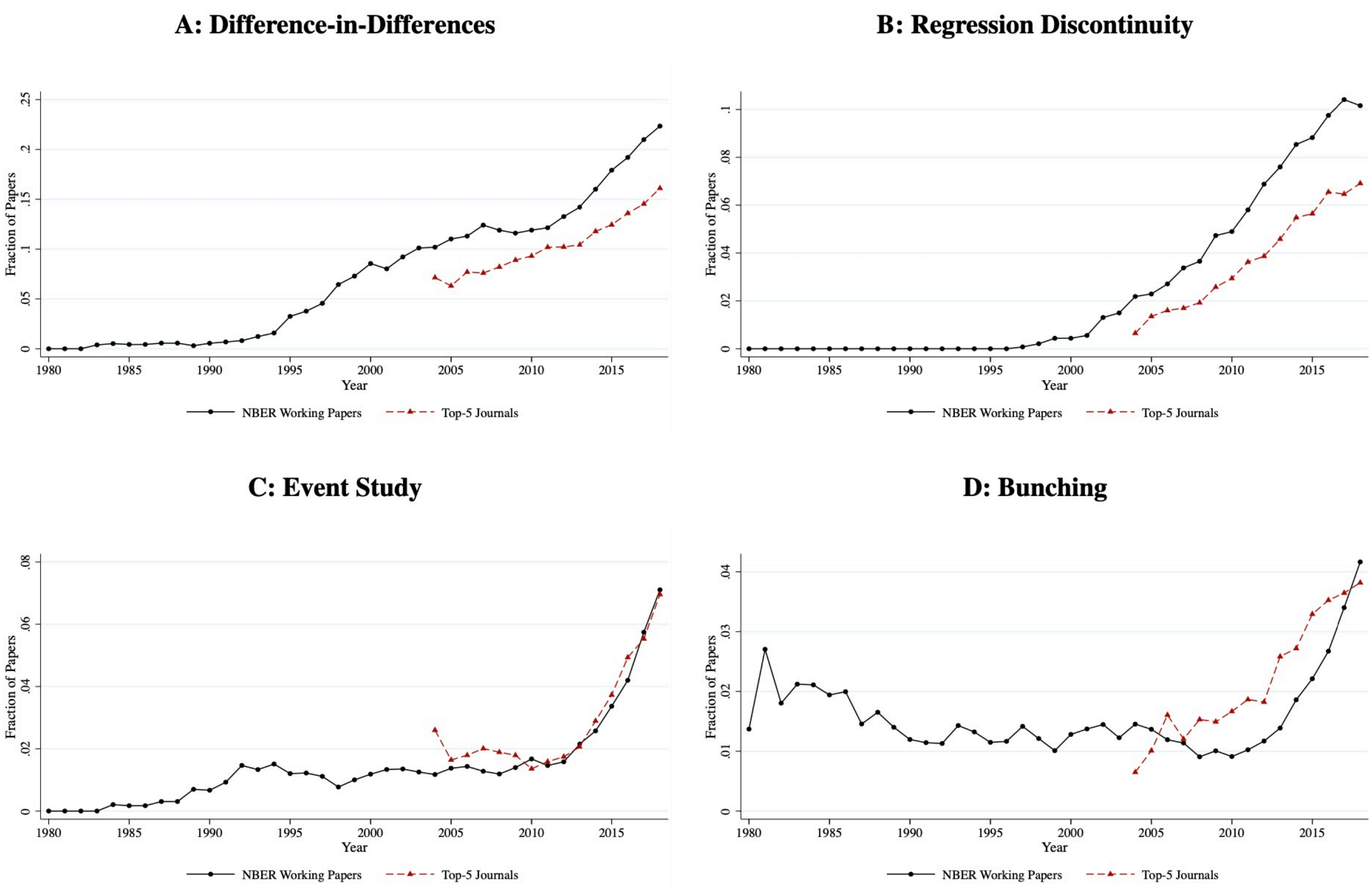
Workshop outline

- Three types of solutions
  - → Aggregated group-time ATT
  - → Stacked regression
  - → Explicit Imputation
- Continuous treatments
- Fuzzy difference-in-differences

What is difference-in-differences (DiD)

- DiD is a very old, relatively straightforward, intuitive research design
- A group of units are assigned some treatment and then compared to a group of units that weren't
- Early usage in several 19th century health policy debates
- Brought into labor economics with Orley Ashenfelter (1978), LaLonde (1986), Card and Krueger (1994)
- Now the most widely used quasi-experimental method

Figure IV: Quasi-Experimental Methods



Notes: This figure shows the fraction of papers referring to each type of quasi-experimental approach. See Table A.I for a list of terms. The series show 5-year moving averages.

Why an entire workshop on DiD?

- **Research advantages**: DiD is often one of the only ways to study large social policies (e.g., decriminalized prostitution [Cunningham and Shah 2018])
- **Worrisome news**: Many new papers suggest canonical methods are biased, maybe severely biased
- **Good news**: Many new solutions and widely available code in both R and Stata
- **Econometrics**: It's always fun to learn econometrics

Pedagogy of the seminar

- Emphasis on assumptions and authors
- It can feel like drinking from a firehose to learn so many papers
- I can’t really advise you on how these are connected to one another, as that level of depth I’m still working on myself

Potential outcomes review

- DiD really can't be understood without committing to some common causality language
- Standard language is the potential outcomes model, sometimes called the Rubin-Neyman model
- Don't go over potential outcomes too fast or you'll miss all the fun
- Potential outcomes are thought experiments about worlds that never existed, but which *could have*

Introduction to Counterfactuals and Causality

- Aliens come and orbit earth, see sick people in hospitals and conclude “these ‘hospitals’ are hurting people”
- Motivated by anger and compassion, they kill the doctors to save the patients
- Sounds stupid, but earthlings do this too - all the time
- Let’s look at the challenges of making causality synonymous with correlations

#1: Correlation and causality are very different concepts

These are not the same thing:

- Causal question: “If a doctor puts a person with Covid on a ventilator (D), will her health (Y) improve?”
- Correlation question:

$$\frac{Cov(D,Y)}{\sqrt{Var_D}\sqrt{Var_Y}}$$

#2: Coming first may not mean causality!

- Every morning the rooster crows and then the sun rises
- If the feral cat had killed the rooster the sun would have still risen, so coming first must not be enough
- *Post hoc ergo propter hoc*: “after this, therefore, because of this”

#3: No correlation does not mean no causality!

- A sailor sails her sailboat across a lake
- Wind blows, and she perfectly counters by turning the rudder
- The same aliens observe from space and say “Look at the way she’s moving that rudder back and forth but going in a straight line. That rudder is broken.” So they send her a new rudder
- They’re wrong but why are they wrong? There is, after all, no correlation
- Question: What if she had been moving the rudder by flipping coins?

Potential outcomes notation

- Let the treatment be a binary variable:

$$D_{i,t} = \begin{cases} 1 \text{ if hospitalized at time } t \\ 0 \text{ if not hospitalized at time } t \end{cases}$$

  where i indexes an individual observation, such as a person
- Potential outcomes:

$$Y_{i,t}^{j} = \begin{cases} 1 \text{ health if hospitalized at time } t \\ 0 \text{ health if not hospitalized at time } t \end{cases}$$

  where j indexes a counterfactual state of the world
- I'll drop t subscript, but note – these are potential outcomes for the same person at the exact same moment in time

Moving between worlds

- A potential outcome Y^1 and a historical outcome Y are neither conceptually nor notationally the same thing
- Potential outcomes are *hypothetical* possibilities describing states of the world but historical outcomes actually occurred
- We choose among potential outcomes by selecting the treatment

Important definitions

Definition 1: Individual treatment effect

The individual treatment effect, δ_i, equals $Y_i^1 - Y_i^0$

Important definitions

Definition 2: Average treatment effect (ATE)

The average treatment effect is the population average of all i individual treatment effects

$$\begin{aligned} E[\delta_i] &= E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1] - E[Y_i^0] \end{aligned}$$

Important definitions

Definition 3: Switching equation

An individual's observed health outcomes, Y, is determined by treatment assignment, D_i, and corresponding potential outcomes:

$$
\begin{aligned}
Y_i &= D_i Y_i^1 + (1 - D_i) Y_i^0 \\
Y_i &= \begin{cases} Y_i^1 \text{ if } D_i = 1 \\ Y_i^0 \text{ if } D_i = 0 \end{cases}
\end{aligned}
$$

So what's the problem?

Definition 4: Fundamental problem of causal inference

If you need both potential outcomes to know causality with certainty, then since it is impossible to observe both Y_i^1 and Y_i^0 for the same individual, δ_i, is *unknowable*.

Conditional Average Treatment Effects

Definition 5: Average Treatment Effect on the Treated (ATT)

The average treatment effect on the treatment group is equal to the average treatment effect conditional on being a treatment group member:

$$\begin{aligned} E[\delta|D=1] &= E[Y^1 - Y^0|D=1] \\ &= E[Y^1|D=1] - E[Y^0|D=1] \end{aligned}$$

Conditional Average Treatment Effects

Definition 6: Average Treatment Effect on the Untreated (ATU)

The average treatment effect on the untreated group is equal to the average treatment effect conditional on being untreated:

$$\begin{aligned} E[\delta|D=0] &= E[Y^1 - Y^0|D=0] \\ &= E[Y^1|D=0] - E[Y^0|D=0] \end{aligned}$$