Difference-in-Differences

ECON 490

Email: tmackay@fullerton.edu

Slides Overview

In these slides, we'll:

- Introduce difference-in-differences (DiD) as ***research design*** and tool for causal inference
- Discuss DiD and event studies in research

Research Designs and Estimators

Research designs are ***conceptual*** approaches used to identify a relationship of interest

- Causal models like DiD, RD, etc. are research designs
- They help us identify the causal effect of X on Y

To actually ***implement*** a research design, we need to pick a particular ***estimator***

- The specific approach we use in R to estimate the relationship
- E.g., `lm()` in R (and OLS more generally) is an ***estimator***

Two Ways of Teaching and Thinking About DiD

"Classic" DiD lecture starts with a 2x2 example

- Two groups, two time periods, one gets a treatment, one doesn't
- Helpful for intuition … but it's a special case

We can also think of DiD as a natural extension of our fixed effects discussion

- Fixed effect for groups + fixed effect for time
- Two-Way Fixed Effects (TWFE) ≈ DiD without a causal interpretation

Two-Way Fixed Effects

Consider the effect of minimum wages on employment

- Different states have different labor markets
- Economic trends vary across the United States

Solution? Control for state FEs + year FEs

- Solves our OVB problem ***IFF*** points above are the only OVs
- Residual variation in MW = idiosyncratic changes to states’ MW over time

Minimum Wages and Employment

Does our TWFE regression ***identify*** the causal effect of MW?

Are we sure every possible OV is either:

1. Reflected in constant avg. differences across states (absorbed by state FE)?
2. Reflected in national trends over time (absorbed by year FE)?

Could something be happening around the same time that states raise MW?

- Correlated in time with employment and MW changes = OVB
- Generates ***selection*** into treatment vs. control

2x2 Difference-in-Differences

Suppose we've got data for OC and LA counties

- Total new homes constructed in each county in 2021 and 2022
- Suppose LA implements a policy to increase housing construction in Jan. 2022

County	Year	New Housing (NH)	Policy
Los Angeles	2021	120	0
Los Angeles	2022	130	1
Orange County	2021	90	0
Orange County	2022	95	0

What's the Effect of LA's Housing Policy?

One potential answer = just compare housing in LA before and after:

$$NH_{LA}^{2022} - NH_{LA}^{2021} = 130 - 120 = 10$$

In words, attribute all new construction in 2022 in LA to this policy

- Does this make sense?
- What if construction generally goes up over time?

A Better Solution

We're worried about there being a trend in housing production

- We've got data on new housing in OC as well!
- Use OC data to identify overall housing trend

Subtracting out the increase in housing in OC let's us remove this trend:

$$\left(NH_{LA}^{2022} - NH_{LA}^{2021}\right) - \left(NH_{OC}^{2022} - NH_{OC}^{2021}\right) = (130 - 120) - (95 - 90) = 5$$

$$\boxed{= \Delta NH_{LA} - \Delta NH_{OC}}$$

Notice that by comparing Δ's, difference in avg. NH across OC and LA drops out

Putting Our Solution in Regression Terms

Let's introduce some quick terminology:

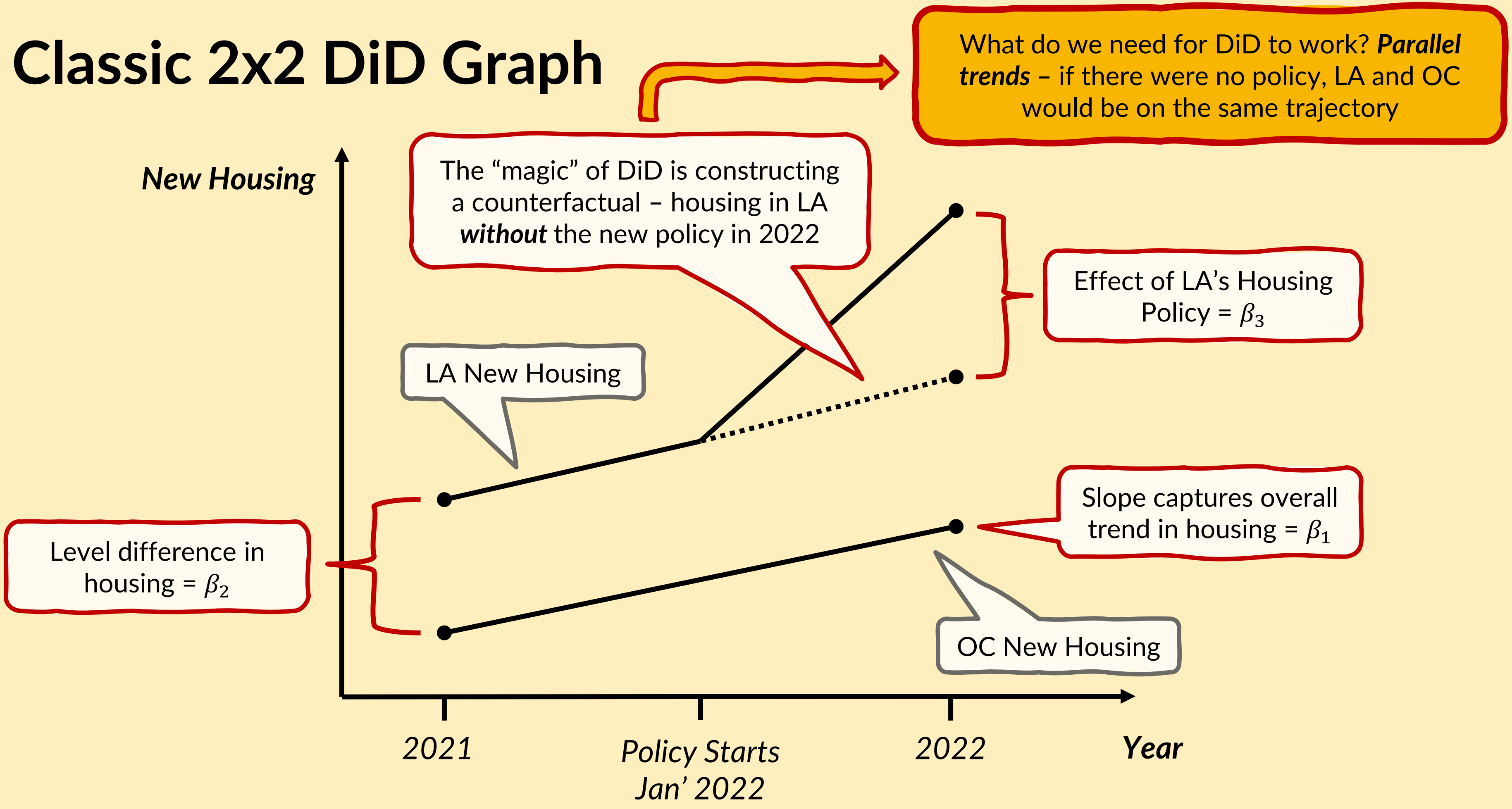
- **Pre-period** is 2021 (year before treatment)
- **Post-period** is 2022 (year after treatment)
- **Treated** county is LA (**control** county is OC)

$$NH_c^t = \beta_0 + \beta_1 Post_t + \beta_2 Treated_c + \beta_3 Post_t \times Treated_c + u$$

DiD estimate of the effect of this policy is β_3

- Change in NH in LA from 2021 to 2022 controlling for overall trend in NH

Classic 2x2 DiD Graph



Another Way of Writing Our Model

How does this connect with TWFE? Consider the following:

$$NH_c^t = \alpha_0 + \alpha_1 Policy_{ct} + \gamma_c + \tau_t + u_{ct}$$

This should look like a normal FE regression equation!

- With 2x2 example, FEs are dummy variables for LA and 2022 respectively
- $Policy_{ct}$ = 1 for having the policy (0 otherwise)

Interpretation of α_1 is the **same** as β_3 from last slide

An Important Distinction

While discussing FEs in metrics review, we said we can always include FEs in a regression

- That doesn't mean we can automatically interpret that regression causally!
- There still might be other sources of OVB

We've highlighted conceptual links between TWFE and DiD to build on prior FE discussion

- In practice, TWFE (via OLS) is an ***estimator*** for DiD ***research designs***

Key distinction – to interpret TWFE regression as DiD model that gives us the causal effect of an explanatory variable, we need to assess the parallel trends assumption

- ***Not*** having parallel trends = we have OVB = we can't make causal claims

DiD in Contemporary Research

Most DiD papers have lots of places getting treated at different times

- Changes the interpretation (and causes estimation problems!)
- Parallel trends assumption becomes “no differential pre-trends”

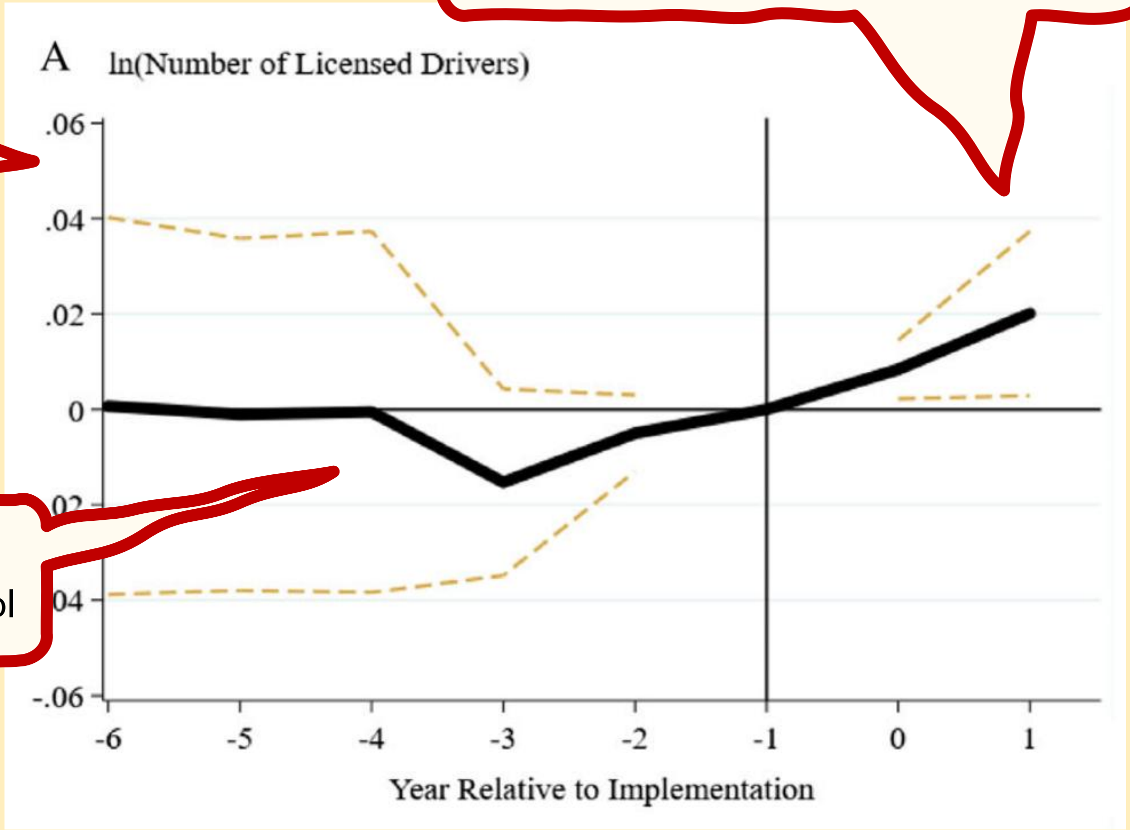
Before treatment, is outcome trending differently for treated vs untreated?

- Generally, you’ll see an ***event study*** graph addressing this
- Zero differences in the pre-treatment time periods are a good thing
- Impact of policy is then the differences in the post-treatment periods

DiD Event Studies

Effect of the policy is post-treatment difference b/w treated and control

Coefficients on dummy variables for years pre- and post policy for treatment states



Relatively flat line pre-treatment = no difference between treatment and control

Event Studies as a Research Design

Sometimes, you'll see papers use event studies without DiD

- Why choose one vs. the other?
- Context matters (finance applications, etc.)
- In general, you need more data for convincing event studies

Same general principles and intuition apply in both cases

- DiD says, "let's see what happened post-treatment ***on average***"
- With event studies, try to see ***dynamic*** effects of treatment over time

Reading Modern DiD Research

Remember that difference-in-differences is a research design

- Lots of ways to ***estimate*** DiD
- Traditionally, synonymous with TWFE + OLS

In the past several years, lots of advances in DiD and event study ***estimators***

- ***Don't*** get bogged down with technical details of estimation!
- Focus on the conceptual framework – carries over from old to new DiD