Difference-in-Differences

Introduction

- Difference-in-differences (DiD) is another tool we have at our disposal to identify causal effects.
- It is similar to using panel data, but is not quite the same. While panel data requires observations of individuals over time, DiD can use observations of larger cohorts over time, e.g. states or countries.
- DiD is used to estimate the causal effect of some binary treatment. Typically, this is some government policy, like a change in the minimum wage or a change to the healthcare system, etc.
- To help understand what DiD is trying to do, consider the following two bad experiments:

Bad Experiments

- **Before** and **after** comparison:
  - **Compares**: same people/units before and after treatment.
  - **Drawback**: doesn't control for time trends.
- **Treated** and **non-treated** comparison:
  - **Compares**: those who are treated with those who are not.
  - **Drawback**: selection bias is ignored - did those who were treated, choose to be treated?

Two Wrongs Make a Right

- The idea of DiD is to combine these two bad experiments to create one good experiment.
- The DiD estimator is essentially:

$$DiD = \left(\bar{Y}^{treated}_{after} - \bar{Y}^{treated}_{before}\right) - \left(\bar{Y}^{non-treated}_{after} - \bar{Y}^{non-treated}_{before}\right).$$

- It is easy to see why this is called the difference-in-differences estimator.
- But to understand why this works, I think it is easiest to see an example.

Minimum Wage Example (Card and Kreuger, 1994)

- Suppose we are interested in the effect of increasing the minimum wage on employment; an important and hotly debated question.
- Some suggest that by increasing the minimum wage, you help out the poorest people because this must surely raise their wage.
- However, this could be a naive interpretation. By increasing the minimum wage, employers may reduce the number of people they hire resulting in higher unemployment and more poverty.
- Card and Krueger (1994) use a dramatic change in the New Jersey state minimum wage to see if this is true. In particular, on April 1st 1992, New Jersey raised the state minimum wage from \$4.25 to \$5.05. (Almost a 20% increase!)

Minimum Wage Example (Card and Kreuger, 1994)

- They collected data on employment at fast-food 'restaurants' (employs a lot of minimum wage people) in February 1992 and again in November 1992.
- They also collected data on the same types of restaurants in Eastern Pennsylvania, just a short distance across the Delaware river. The minimum wage here stayed constant at $4.25.
- The DiD analysis basically just compares the change in employment in New Jersey to the change in employment in Pennsylvania around this time.
- We will use the 'potential outcomes' notation we just learnt about to see how this works...

Potential Outcomes for DiD

- DiD can be thought of as an aggregate data version of fixed-effects. Let

$$\begin{aligned} Y_{1ist} &= \text{employment at restaurant i in state s} \\ &\quad \text{in period t if there is a HIGH minimum wage.} \\ Y_{0ist} &= \text{employment at restaurant i in state s} \\ &\quad \text{in period t if there is a LOW minimum wage.} \end{aligned}$$

- These are potential outcomes - we only get to see one or the other for each i, s, t. For example, we see Y_{1ist} in NJ in November but we don't see Y_{0ist} then.
- We assume the following additive structure

$$E\left[Y_{0ist} \middle| s, t\right] = \gamma_s + \lambda_t.$$

- This says that without any changes in the minimum wage, employment is determined as the sum of a state effect and a time effect. (See the similarity to panel data!)

Potential Outcomes for DiD

- Now, let D_{st} be a dummy equal to 1 if state s has a high minimum wage in period t.
- If we assume that the treatment effect is constant, i.e. $E\left[Y_{1ist} - Y_{0ist} \middle| s, t\right] = \beta$, we can write

$$E\left[Y_{ist} \middle| s, t\right] = \gamma_s + \lambda_t + \beta D_{st}$$

or, in regression form

$$Y_{ist} = \gamma_s + \lambda_t + \beta D_{st} + \epsilon_{ist}.$$

- Consider the following differences

$$E\left[Y_{ist} \middle| s = PA, t = Nov\right] - E\left[Y_{ist} \middle| s = PA, t = Feb\right] = \lambda_{Nov} - \lambda_{Feb}$$

and

$$E\left[Y_{ist} \middle| s = NJ, t = Nov\right] - E\left[Y_{ist} \middle| s = NJ, t = Feb\right] = \lambda_{Nov} - \lambda_{Feb} + \beta.$$

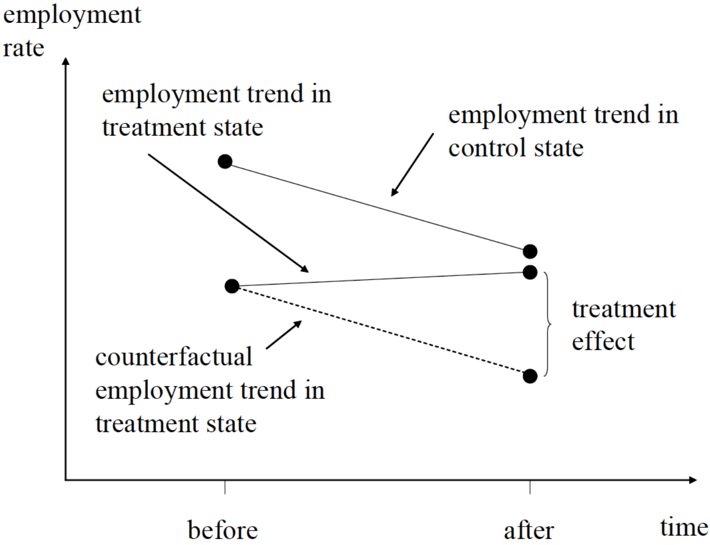
DiD Equation

- Combining these last two equations... in particular, taking the **difference of these differences**, we get

$$
\begin{aligned}
E\left[Y_{ist}\middle|s=PA,t=Nov\right]-E\left[Y_{ist}\middle|s=PA,t=Feb\right]& \\
-\left(E\left[Y_{ist}\middle|s=NJ,t=Nov\right]-E\left[Y_{ist}\middle|s=NJ,t=Feb\right]\right) &= \beta
\end{aligned}
$$

- We have managed to isolate the causal effect of a change in the minimum wage!
- The first difference removes the state effects, and the second difference removes the time effect. So what we are left with is the treatment effect of the change in the minimum wage.
- We can see what is going on graphically...

Graphical Representation of DiD



(Credit: Mostly Harmless Econometrics, Angrist and Pischke)

Card and Kreuger (1994) Results

- So, what do the results show...

Variable		PA (i)	NJ (ii)	Difference, NJ-PA (iii)
1.	FTE employment before,	23.33	20.44	-2.89
	all available observations	(1.35)	(0.51)	(1.44)
2.	FTE employment after,	21.17	21.03	-0.14
	all available observations	(0.94)	(0.52)	(1.07)
3.	Change in mean FTE	-2.16	0.59	2.76
	employment	(1.25)	(0.54)	(1.36)

(Credit: Mostly Harmless Econometrics, Angrist and Pischke)

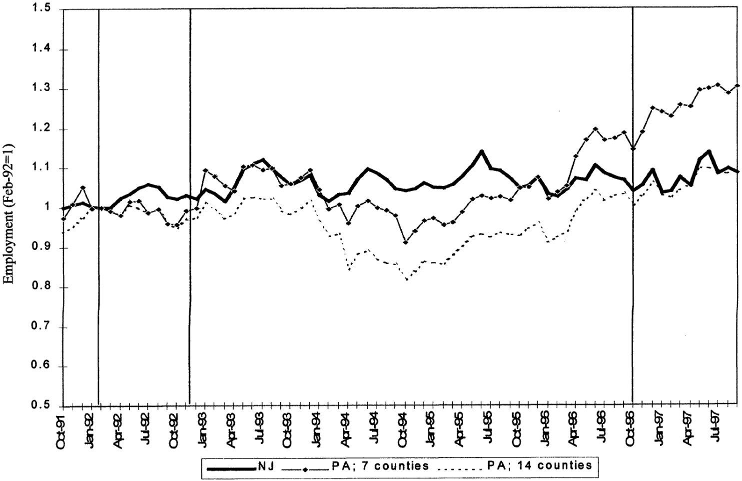
- The bottom right number, 2.76, with its corresponding standard error, 1.36, show a **positive effect** on employment of an increase in the minimum wage... this is the opposite of what we expected!

- How convinced are we of these results?

Checking the Identification Assumption

- The key identifying assumption here is that employment trends would be the same in both states in the absence of treatment, so that treatment induces a deviation from this common trend.
- We can investigate whether this assumption holds by using data on multiple periods. In fact, this is exactly what Card and Kreuger do in a follow-up study in 2000...
- On the next slide we see a time series of employment for each state. The first two vertical lines indicate Feb and Nov 1992, the third line shows when PA raised their minimum wage to $4.75.

Checking the Identification Assumption



(Credit: Mostly Harmless Econometrics, Angrist and Pischke)

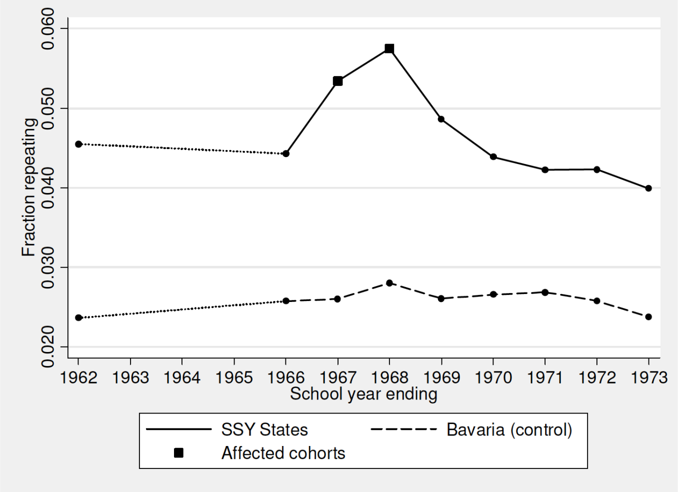
Checking the Identification Assumption

- The data reveals fairly substantial year-to-year employment variation. These swings often seem to differ substantially in the two states.
- So Pennsylvania may not provide a very good measure of counterfactual employment rates in New Jersey in the absence of a policy change, and vice versa.
- The assumption that the employment in each state, in the absence of a policy change, is just the sum of a time trend and a state effect, does not seem to hold.
- We cannot take much from these results then :(

School Term Length Example (Pischke, 2007)

- Here is a better example...
- Until the 1960s, children in all German states except Bavaria started school in the Spring. Beginning in the 1966-67 school year, the Spring-starters moved to start school in the Fall. The transition to a Fall start required two short school years for affected cohorts, 24 weeks long instead of 37.
- So, students in these cohorts effectively had their time in school compressed relative to cohorts on either side and relative to students in Bavaria.
- The graph on the next slide plots the likelihood of grade repetition for the each of the 1962-73 cohorts of 2nd graders in Bavaria and affected states.

School Term Length Example (Pischke, 2007)



(Credit: Mostly Harmless Econometrics, Angrist and Pischke)

School Term Length Example (Pischke, 2007)

- Repetition rates in Bavaria were reasonably flat from 1966 onwards at about 2.5%. While repetition rates are higher in the short-school-year states, at around 4.5% in 1962 and 1966, before the change in term length.

- However, repetition rates jump up by about a percentage point for the two affected cohorts in these states, before falling back to the baseline level. This graph provides strong visual evidence of treatment and control states with a common time trend, and a treatment effect that induces a sharp but temporary deviation from this trend.

- So, a shorter school year seems to have increased repetition rates for affected cohorts. And we can be reasonably confident in this effect.

Regression DiD

- Although we can calculate the DiD estimator by simply calculating the mean outcome for each of the four situations (Pre-Treatment Control, Post-Treatment Control, Pre-Treatment Treated, Post-Treatment Treated), it might be nicer to use regression.
- Using regression allows us to automatically get a standard error for our causal effect, and it also allows us to add in control variables if we like.
- In the minimum wage example, the regression model includes a dummy for the state, a dummy for the year, and an interaction term made up of these two dummies.

$$Y_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \beta \left(NJ_s \times d_t \right) + \delta X_{ist} + \epsilon_{ist}$$

where NJ_s is a dummy for New Jersey, d_t is a dummy for November, and X_{ist} are an optional set of control variables.

Regression DiD - Treatment Intensity

- A third advantage of using regression for DiD is the ability to allow for **different treatment intensities**.
- Instead of New Jersey and Pennsylvania in 1992, for example, we might look at all state minimum wages in the United States. Some of these are a little higher than the federal minimum (which covers everyone regardless of where they live), some are a lot higher, and some are the same.
- The minimum wage is therefore a variable with differing intensity.
- In addition to statutory variation in state minima, the local importance of a minimum wage varies with average state wage levels. For example, the early-1990s Federal minimum of $4.25 was probably irrelevant in Connecticut - since they have high average wages - but a big deal in Mississippi.

Minimum Wage Example (Card, 1992) - Treatment Intensity

- In 1990, the federal min. wage increased from \$3.35 to \$3.80.
- Card (1992) exploits regional variation in the impact of the federal minimum wage to look at how this change affected wages of teens. The model looks like this

$$Y_{ist} = \gamma_s + \lambda_t + \beta\left(FA_s \times d_t\right) + \epsilon_{ist}$$

  where FA_s aims to measure the **fraction affected** by the minimum wage (it actually measures the proportion of teens in the labour force earning less than \$3.80, before the increase in minimum wage), and d_t is a dummy equal to 1 for observations after 1990.

Minimum Wage Example (Card, 1992) - Treatment Intensity

- As in the NJ/PA study, Card (1992) works with data before and after, in this case 1989 and 1992. But this study uses all states.
- In contrast to the NJ/PA study where the interaction term was 'turned off or turned on', the interaction term now is 'turned off or turned on to some degree'.
- Card actually first differenced the regression and worked with state averages before estimating it, i.e. the model is

$$\Delta \bar{Y}_s = \lambda^* + \beta FA_s + \Delta \bar{\epsilon}_s$$

Card (1992) Results

Explanatory Variable		Equations for Change in Mean Log Wage: (1)	(2)	Equations for change in Teen Employment-Population Ratio: (3)	(4)
1.	Fraction of Affected Teens	0.15 (0.03)	.14 (0.04)	0.02 (0.03)	-.01 (0.03)
2.	Change in Overall Emp./Pop. Ratio	–	0.46 (0.60)	–	1.24 (0.60)
3.	R-squared	0.30	0.31	0.01	0.09

(Credit: Mostly Harmless Econometrics, Angrist and Pischke)

- We concentrate on columns (1) and (3) (the other two just have an extra control variable).
- Column (1) shows that wages increased more in states where the fraction of affected teens is higher.
- Column (3) shows that the effect on employment is unrelated to the fraction affected. (Similar to the finding in the NJ/PA paper).

DiD in R

- I'm not going to cover how to do DiD in R here.
- Problem set 8 covers this in full, but it is not really any different to what we have done before.

Some Recent Examples

- To finish, I wanted to mention a couple of recent papers that use DiD in interesting settings.
- Grogger and Ridgeway (2006) use the daylight saving time shift to develop a police racial profiling test that is based on differences in driver race visibility and (hence) the race distribution of traffic stops across daylight and darkness. They found no discrimination.

Some Recent Examples

- However, urban areas may be well-lit at night, eroding the power of their test.
- Horace and Rohlin (2016) refine their test using satellite streetlight location data. The results change in the direction of finding profiling of black drivers. They suggest that the odds of a black driver being stopped (relative to nonblack drivers) increase 15% in daylight compared to darkness.

Some Recent Examples

- Price and Wolfers (2010) look at whether NBA referees display racial preferences. Their DiD analysis is given below (let's concentrate on the extreme cases).

TABLE III

DIFFERENCES IN DIFFERENCES: FOUL RATE (=48 × FOULS/MINUTES PLAYED)

	Black players	White players	Difference: black–white foul rate	Slope: Δ(black–white)/ Δ%white refs
0% white refs	4.418	5.245	−0.827	
(n = 7,359)	(0.043)	(0.094)	(0.106)	
33% white ref	4.317	4.992	−0.675	0.455
(n = 54,537)	(0.016)	(0.035)	(0.038)	(0.331)
67% white refs	4.335	4.989	−0.654	0.064
(n = 126,317)	(0.010)	(0.023)	(0.025)	(0.137)
100% white refs	4.322	4.897	−0.574	0.240**
(n = 78,771)	(0.013)	(0.029)	(0.032)	(0.121)
				Diff-in-diff
Average slope:	−0.022	−0.204***		0.182***
Δfouls/Δ%white refs	(0.027)	(0.066)		(0.066)
				(n = 906)

Some Recent Examples

- I think their analysis is a little incorrect...
- Write out the four estimates as being made up of a 'black player effect', a 'black referee effect', and a 'same race effect'. Then look at the difference-in-differences calculation needed to obtain only the 'same race effect'.
- Then our DiD is calculated as
  $[(4.897 - 5.245) - (4.322 - 4.418)]/2 = -0.126$
- So, the **difference in the differences** comes out at -0.126, suggesting that a white player earns 0.13 fewer fouls (per 48 minutes played) when facing three white referees than when facing three black referees.

Summary

- We have seen the difference in differences approach to identifying causal effects.
- Through a couple of examples we showed what assumptions we need to get a causal interpretation, namely, that there is a constant trend in both groups.
- A third example showed how we can allow for the intensity of the treatment to vary across groups.