

Big Data and Economics

Difference-in-differences

Kyle Coombs

Bates College | DCS/ECON 368

Table of contents

- Prologue
- Difference what with who?
 - Effect of time
- Example: Earned Income Tax Credit
- Many groups
- Example: Pandemic Unemployment Assistance

Prologue

Plan for the day

- Computers closed: Work through the intuition of difference-in-differences
- Computers open: Work through practical examples
- By the end of the day you will:
 - Understand how difference-in-differences uses variation in time and treatment to identify causal effects
 - Be able to implement fixed effects models in R
 - Be able to implement difference-in-differences models in R
 - Create pretty event study plots of your results

Check-in on causality

- We've been talking about causality for a bit now
- Popcorn style: what do we need for causality?

Check-in on causality

- We've been talking about causality for a bit now
- Popcorn style: what do we need for causality?
- We need to account for everything that is correlated with treatment
- There are two extreme strategies:
 1. Control variables for everything!
 - Did we control for it all? How do we know?
 2. Get rid of everything correlated with treatment!
 - How do we know we got rid of all the correlation?

And then there's a middle ground...

Bridge from fixed effects

- Fixed effects let us control for everything that varies between groups
 - If we have fixed effects for an individual, we account for how much one individual differs from another **on average**
 - Effectively, we control for their hometown, family, education, race, etc. without including those variables in the model
- If treatment is randomly assigned within the fixed effect, we're done!
 - We've controlled for everything correlated with the treatment
 - Pro: control for a bunch of stuff
 - Con: washes out a lot of variation! Result can be noisier if there's not much within-variation to work with

What if things change over time?

- Causal identification with fixed effects requires no endogenous variation within groups over time
 - If there is, we need to do something else
- What does "endogenous" variation mean?

What if things change over time?

- Causal identification with fixed effects requires no endogenous variation within groups over time
 - If there is, we need to do something else
- What does "endogenous" variation mean?
- That might be a tricky assumption!

Well what if something changes?

- Let's say something within fixed effects changes over time
 - e.g. The government increases the Earned Income Tax Credit (EITC) for a group of people who start to work less
 - Does that mean the increased EITC caused them to work less?

Well what if something changes?

- Let's say something within fixed effects changes over time
 - e.g. The government increases the Earned Income Tax Credit (EITC) for a group of people who start to work less
 - Does that mean the increased EITC caused them to work less?
 - No! We don't know what would have happened if the EITC wasn't increased
 - Maybe this group was already unable to work and that's why the EITC was increased

What about that other group?

- But what about people who didn't get the EITC increase?
 - We can compare them to the people who did get the EITC increase
- Can we compare them after the EITC increase? No, because they're different groups
 - Employment levels might have always differed between these two groups
- So we need to account for variation between the groups **and** time
 - This is the idea behind difference-in-differences

Difference what with who?

Difference-in-differences

- Today we will talk about difference-in-differences (DiD), which is a way of using within variation in a more deliberate way in order to identify the effect we want
- All we need is a treatment that *goes into effect* at a particular time, and we need a group that is *treated* and a group that is *not*
- Then, we compare the within-variation for the treated group vs. the within-variation for the untreated group
- Voila, we have an effect!

Difference-in-Differences

- Because the requirements to use it are so low, DiD is used a *lot*
- Any time a policy is enacted but isn't enacted everywhere at once? DiD!
- Plus, the logic is pretty straightforward

Difference-in-Differences

- The question DiD tries to answer is "what was the effect of (some policy) on the people who were affected by it?"
- We have some data on the people who were affected both before the policy went into effect and after
- However, we can't just compare before and after, because things change over Time for reasons unrelated to Treatment

Difference-in-Differences

- Why not just control for Time?
 - We can certainly measure that and we can just add a fixed effect
- What happens if we add a fixed effect for Time?

Difference-in-Differences

- Why not just control for Time?
 - We can certainly measure that and we can just add a fixed effect
- What happens if we add a fixed effect for Time?
- It would wash out all the variation!
- You're either Before and Untreated, or After and Treated.
- If you control for Time, you're comparing people with the same values of Time
- But that means they have the same values of Treatment!
- We need to compare Treated and Untreated to get the effect

Difference-in-Difference

- We control for Group by isolating within variation (comparing After to Before)
- Then we control for Time by comparing two sources of Within variation
- We ask *how much more increase* was there in Treatment than control?
- The Control change is *probably* the change you would expect without Treatment
- So anything on *top of that* is the effect of Treatment
- Now we've controlled for Group and Time, identifying the effect!

Extremely bizarre example

- Let's say we have a pill that's supposed to make you taller
- Give it to a kid Adelaide who is 48 inches tall
- Next year they're 54 inches tall - a six inch increase! But they probably would have grown some anyway without the pill. Surely the pill doesn't make you six inches taller.
- SO we compare them to their twin Bella, who started at 47 inches but we DON'T give a pill to
- Next year that twin is 51 inches tall - a four inch increase. So Adelaide probably would have grown about 4 inches without the pill. So the pill boosted her by $(54 - 48) - (51 - 47) = 6 - 4 = 2$ additional inches
- "Adelaide, who was Treated, grew by two *more inches* than Bella, who was Untreated, did over the same period of time, so the pill made Adelaide grow by two inches"
- That's DiD!

Example



Example



Difference-in-Differences

What *changes* are included in each value?

- **Untreated Before**: Untreated Group Mean
- **Untreated After**: Untreated Group Mean + Time Effect
- **Treated Before**: Treated Group Mean
- **Treated After**: Treated Group Mean + Time Effect + **Treatment Effect**
- **Untreated After** - **Untreated Before** = Time Effect
- **Treated After** - **Treated Before** = Time Effect + **Treatment Effect**

$$\text{DiD} = (\text{TA} - \text{TB}) - (\text{UA} - \text{UB}) = \text{Treatment Effect}$$

(Abbreviations for Untreated and Treated Before/After to save space)

Concept Checks

- Why do we need a control group? What does this let us do?
- What do we need to assume is true about our control group?
- In 2015, a new, higher minimum wage went into effect in Seattle, but this increase did not occur in some of the areas surrounding Seattle.
 - How might you use DiD to estimate the effect of this minimum wage change on employment levels?

Difference-in-Difference

- What if there are more than four data points?
- Usually these four points would be four means from lots of observations, not just two people in two time periods
- How can we do this and get things like standard errors, and perhaps include controls?
- Why, use OLS regression of course!

Difference-in-Differences

- We can use what we know about binary variables and interaction terms to get our DiD

$$Y_{it} = \beta_0 + \beta_1 \textit{After}_t + \beta_2 \textit{Treated}_i + \beta_3 \textit{After}_t \textit{Treated}_i + \varepsilon_{it}$$

where \textit{After}_t is a binary variable for being in the post-treatment period, and $\textit{Treated}_i$ is a binary variable for being in the treated group

##	Person	Time	Height	After	Treated
## 1	Adelaide	Before	48	FALSE	TRUE
## 2	Adelaide	After	54	TRUE	TRUE
## 3	Bella	Before	47	FALSE	FALSE
## 4	Bella	After	51	TRUE	FALSE

Difference-in-Differences

- How can we interpret this using what we know?

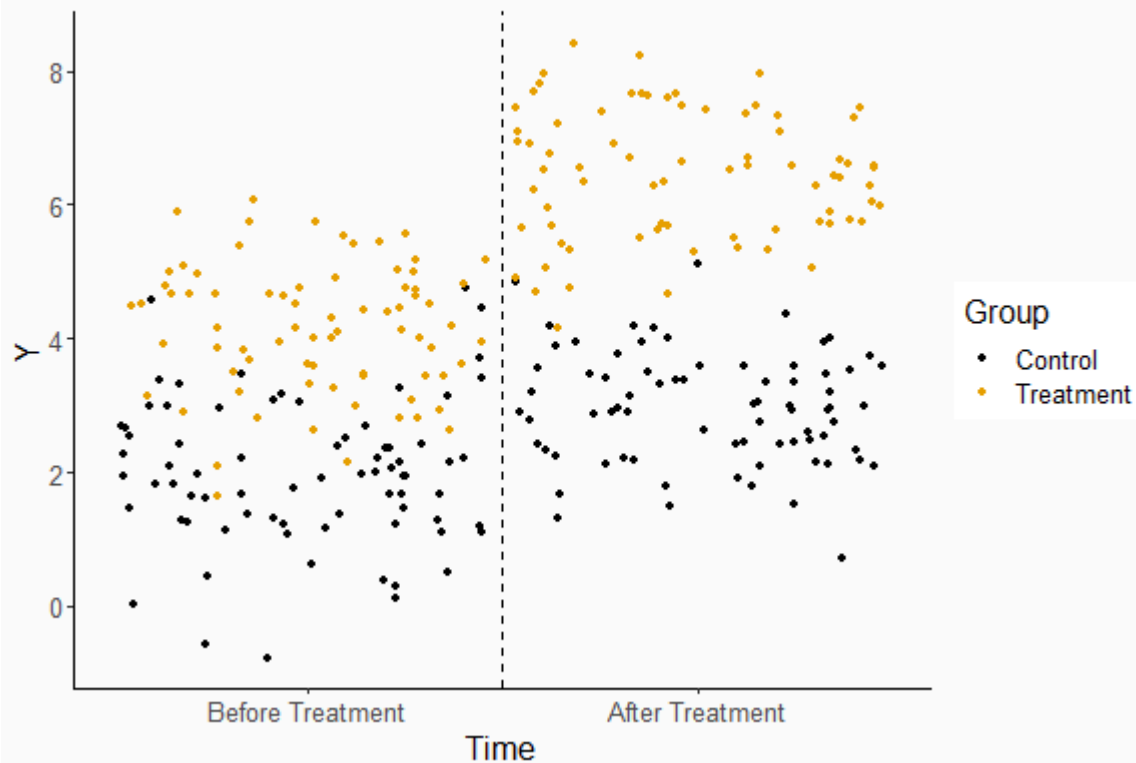
$$Y_{it} = \beta_0 + \beta_1 \textit{After}_t + \beta_2 \textit{Treated}_i + \beta_3 \textit{After}_t \times \textit{Treated}_i + \varepsilon_{it}$$

- β_0 is the prediction when $\textit{Treated}_i = 0$ and $\textit{After}_t = 0 \rightarrow$ the **Untreated Before** mean!
- β_1 is the *difference between* Before and After for $\textit{Treated}_i = 0 \rightarrow$ Untreated (**After - Before**)
- β_2 is the *difference between* Treated and Untreated for $\textit{After}_t = 0 \rightarrow$ **Before (Treated - Untreated)**
- β_3 is *how much bigger the Before-After difference is for* $\textit{Treated}_i = 1$ than for $\textit{Treated}_i = 0 \rightarrow (\text{TA} - \text{TB}) - (\text{UA} - \text{UB}) = \text{DID!}$

Graphically

The Difference-in-Difference Effect of Treatment

1. Start with raw data.



Design vs. Regression

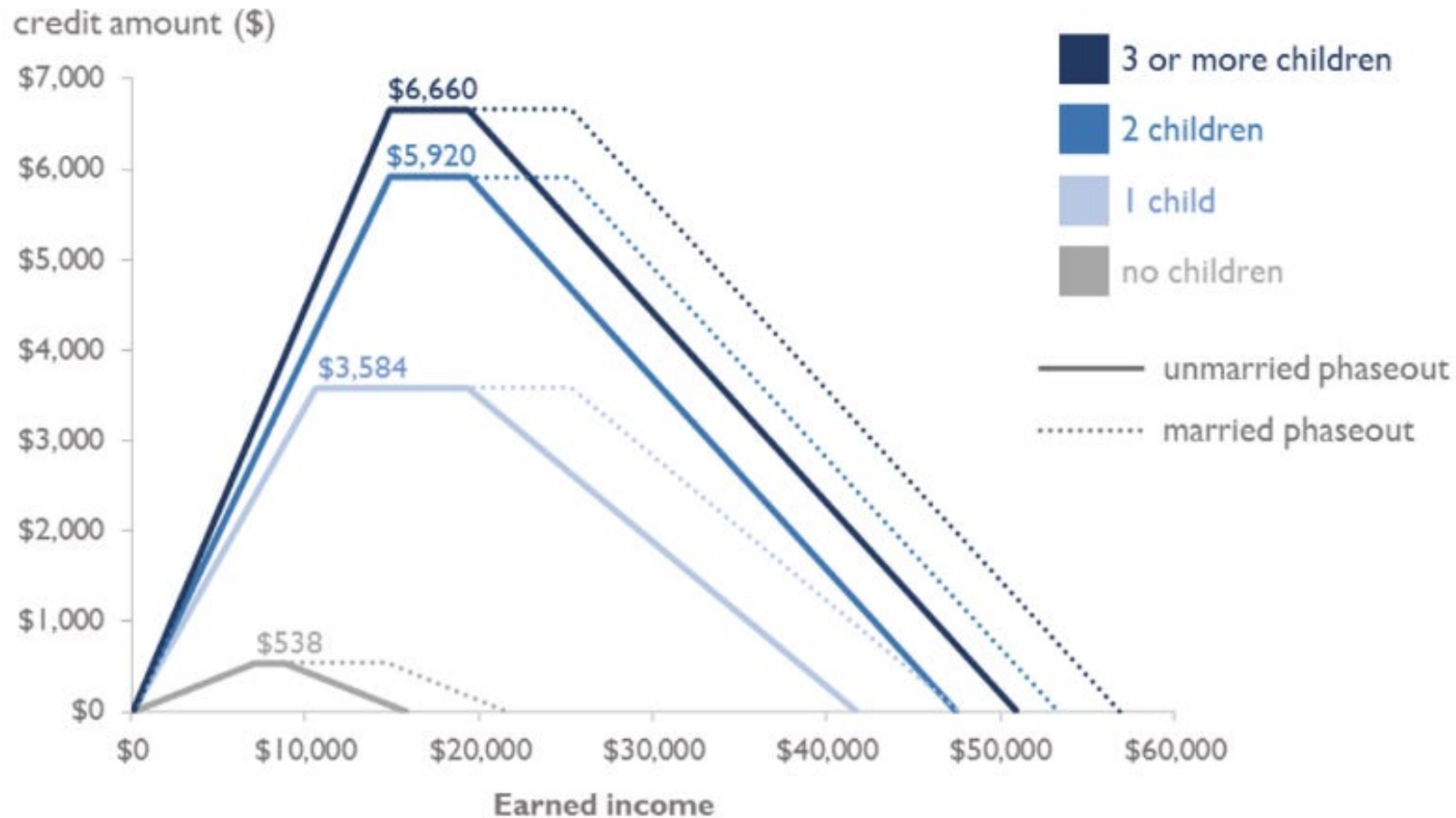
- There is a distinction between *regression model* and *research design*
 - This is also true for fixed effects
- We have a model with an interaction term
- Not all models with interaction terms are DID!
- It's DID because it's an interaction between treated/control and before/after
- If you don't have a before/after, or you don't have a control group, that same setup may tell you something interesting but it won't be DID!

Example: Earned Income Tax Credit

Earned Income Tax Credit

- The Earned Income Tax Credit (EITC) is a refundable tax credit for low-income workers
- Key features:
 - Initially increases as earned income increases before leveling off at a threshold, at next threshold it declines as income increases further
 - Increases with number of children
 - Phaseout differs by filing status (single vs. married)
- It was introduced in 1975 and expanded in 1986, 1990, and 1993 (we'll focus on 1993)
- Policy details vary, so we're gonna focus on single mothers vs. single women without kids to keep it simple
 - Very common to zoom in on groups where we expect the effect to exist
 - Helps you get a sense of what's going on before you try to explain everything

EITC Amount by Number of Qualifying Children, Marital Status, and Income, 2020



Earned Income Tax Credit

- EITC was increased in 1993. This may increase chances single mothers (treated) return to work, but likely not affect single non-moms (control)
- Does this program incentivize single mothers to work more?

```
df <- read.csv('http://nickchk.com/eitc.csv') %>%  
  mutate(after = year ≥ 1994,  
         treated = children > 0)  
df %>%  
  group_by(after, treated) %>%  
  summarize(proportion_working = mean(work))
```

```
## # A tibble: 4 × 3  
## # Groups:   after [2]  
##   after treated proportion_working  
##   <lgl> <lgl>          <dbl>  
## 1 FALSE FALSE          0.575  
## 2 FALSE TRUE           0.446  
## 3 TRUE  FALSE          0.573  
## 4 TRUE  TRUE           0.491
```

Example

- We can do it by just comparing the points, like we did with Adelaide and Bella
- This will give us the DID estimate: The EITC increase increases the probability of working by 4.7 percentage points
- But not standard errors, or the ability to include controls easily

```
means <- df %>%  
  group_by(after, treated) %>%  
  summarize(proportion_working = mean(work)) %>%  
  pull(proportion_working)  
(means[4] - means[2]) - (means[3] - means[1])
```

```
## [1] 0.04687313
```

Let's try OLS!

```
etable(feols(work ~ after*treated, data = df), digits = 3, fitstat=c('n', 'r2'))
```

```
##                                feols(work ~ af..  
## Dependent Var.:                work  
##  
## Constant                      0.575*** (0.009)  
## afterTRUE                     -0.002 (0.013)  
## treatedTRUE                   -0.129*** (0.012)  
## afterTRUE x treatedTRUE       0.047** (0.017)  
## -----  
## S.E. type                      IID  
## Observations                  13,746  
## R2                            0.01260  
## ---  
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Concept Checks

- There are four coefficients. Interpret them carefully in a sentence each.
- Why can the *After* × *Treated* interaction stand in for *Treated*?

What if there are more groups?

- You can recognize $After_t$ and $Treated_i$ as *fixed effects*
- $Treated_i$ is a fixed effect for group - we only need one coefficient for it since there are only two groups
- And $After_t$ is a fixed effect for *time* - one coefficient for two time periods
- You can have more than one set of fixed effects like this! Our interpretation is now within-group *and* within-time
- (i.e. comparing the within-group variation across groups)

Example: Multiple Groups

Multiple Treated and Control Groups

- We can extend DID to having more than two groups, some of which get treated and some of which don't
- And more than two time periods! Multiple before and/or multiple after
- We don't have a full set of interaction terms, we still only need the one, which we can now call *CurrentlyTreated_{it}*
- **If you have more than two groups and/or more than two time periods, then this is what you should be doing**

Multiple treatment groups

- Let's make some quick example data to show this off, with the first treated period being period 7 and the treated groups being 1 and 9, and a true effect of 3

```
did_data <- tibble(group = sort(rep(1:10, 10)),
                    time = rep(1:10, 10)) %>%
  mutate(CurrentlyTreated = group %in% c(1,9) & time ≥ 7) %>%
  mutate(Outcome = group + time + 3*CurrentlyTreated + rnorm(100))
did_data
```

```
## # A tibble: 100 × 4
##   group time CurrentlyTreated Outcome
##   <int> <int> <lgl>          <dbl>
## 1     1     1     FALSE          1.50
## 2     1     2     FALSE          2.43
## 3     1     3     FALSE          3.70
## 4     1     4     FALSE          5.75
## 5     1     5     FALSE          4.57
## 6     1     6     FALSE          6.47
## 7     1     7     TRUE          11.1
## 8     1     8     TRUE          12.6
## 9     1     9     TRUE          12.6
## 10    1    10     TRUE          13.6
## # i 90 more rows
```

Multiple treatment groups

```
# Put group first so the clustering is on group
many_periods_did <- feols(Outcome ~ CurrentlyTreated | group + time, data = did_data)
etable(many_periods_did)
```

```
##                               many_periods_did
## Dependent Var.:                Outcome
##
## CurrentlyTreatedTRUE 3.095*** (0.4991)
## Fixed-Effects:      -----
## group                                Yes
## time                                Yes
## -----
## S.E.: Clustered                by: group
## Observations                    100
## R2                             0.96131
## Within R2                      0.34238
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Two-way fixed effects model

- We just ran a two-way fixed effects model
- We have a fixed effect for group and a fixed effect for time
- This is often done when we have a panel dataset with individuals over time
- The generic formula is:

$$y_{it} = \underbrace{\alpha_i}_{\text{Individual FE}} + \underbrace{\alpha_t}_{\text{Time FE}} + \beta \textit{CurrentlyTreated}_{it} + \varepsilon_{it}$$

Downers and Assumptions

- So... does this all work?
- That example got pretty close to the truth of 3 but who knows in other cases!
- What needs to be true for this to work?

Does it Work?

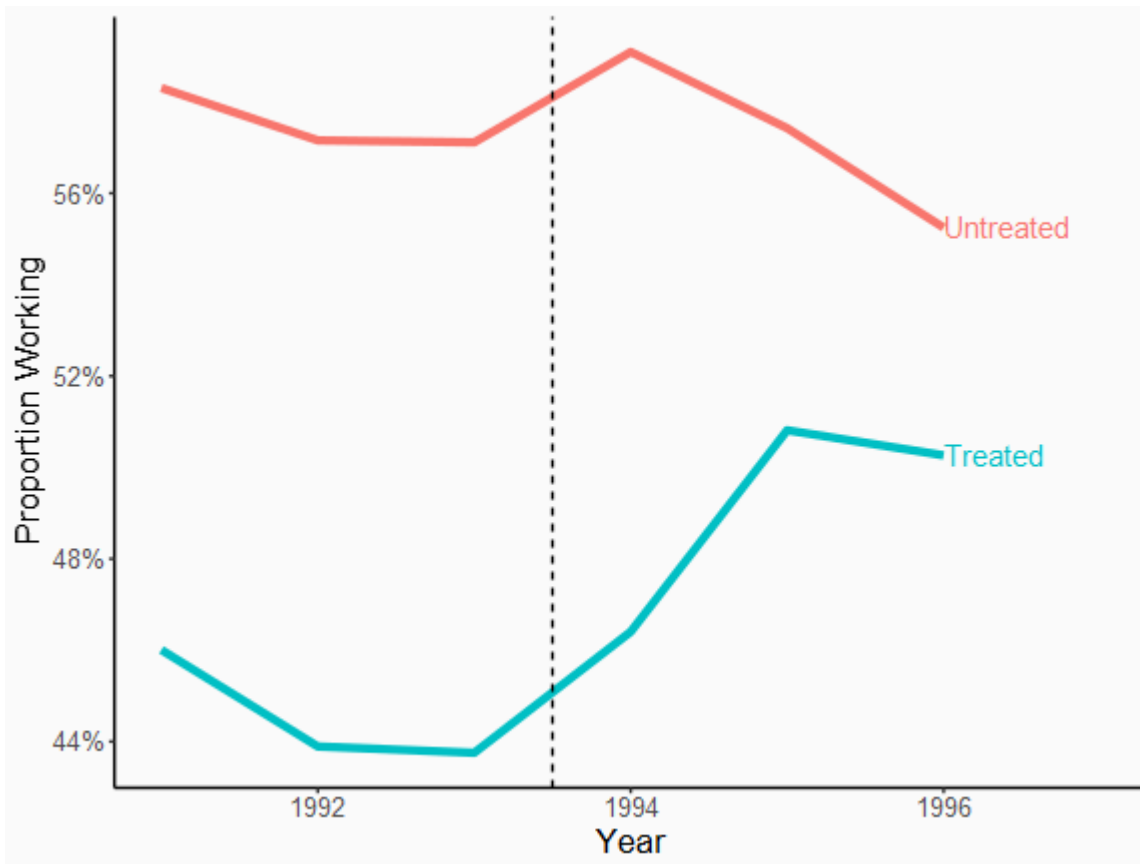
- This gives us a causal effect as long as *the only reason the gap changed* was the treatment
- If Adelaide would have grown six inches *anyway*, then the gap would have grown by 2, pill or not.
 - But we don't know that and mistakenly say it helped her grow by 2
- For fixed effects to have a causal effect with panel data, we assume no uncontrolled endogenous variation across time
- In DID, we need to assume that there's no uncontrolled endogenous variation *across this particular before/after time change*
- An easier assumption to justify but still an assumption!

Parallel Trends

- This assumption - that nothing else changes at the same time, is the poorly-named "parallel trends"
- Again, this assumes that, *if the Treatment hadn't happened to anyone*, the gap between the two would have stayed the same
- Sometimes people check whether this assumption is plausible by seeing if *prior trends* are the same for Treated and Untreated - if we have multiple pre-treatment periods, was the gap changing a lot during that period?
- Sometimes people also "adjust for prior trends" to fix parallel trends violations, or use related methods like Synthetic Control

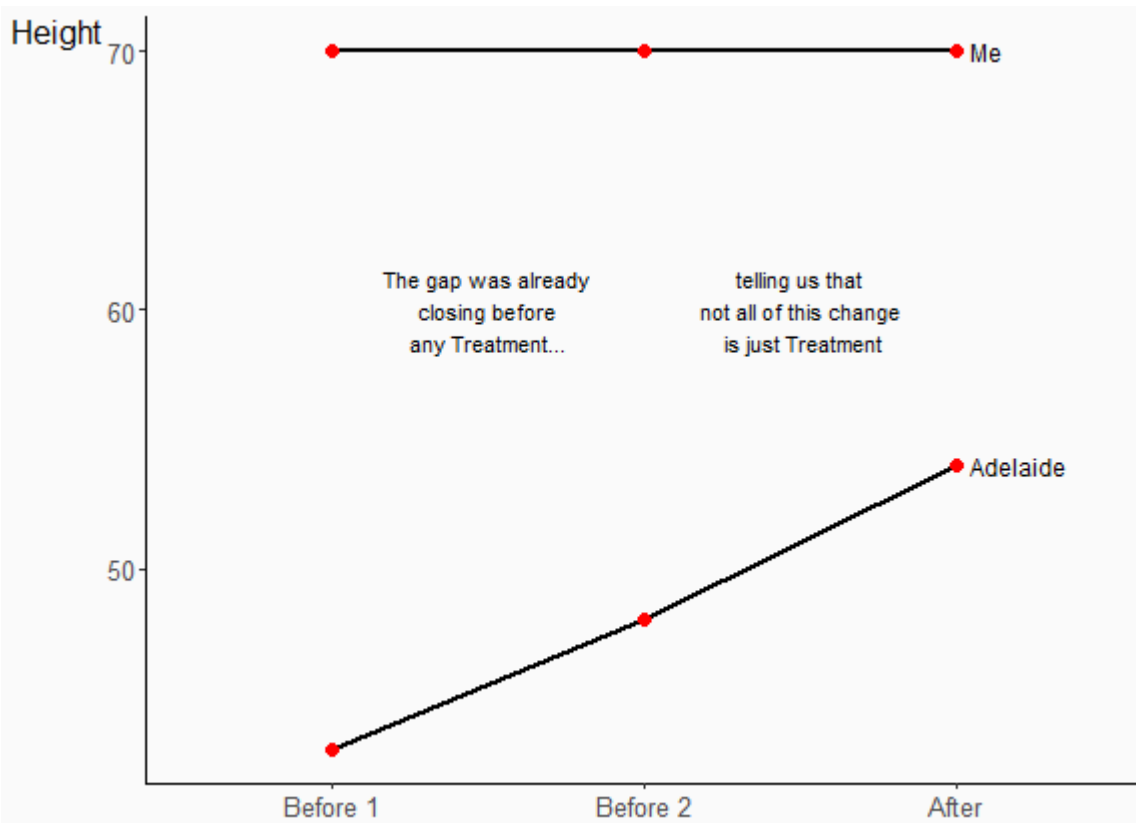
Prior Trends

- Let's see how that EITC example looks in the leadup to 1994
- They look like the gap between them is pretty constant before 1994! They move up and down but the *gap* stays the same. That's good.



Prior Trends

- Formally, prior trends being the same tells us nothing about parallel trends
- But it can be suggestive.
 - Going back to the height pill example, what if instead of comparing Adelaide and Bella, child twins, we compared Adelaide to *me*?
- Seeing the gap close *anyway* in previous years is a clue: it's not just the pill



Parallel Trends

- Just because *prior trends* are equal doesn't mean that *parallel trends* holds.
- *Parallel trends* is about what the before-after change *would have been* - we can't see that!

Parallel Trends

- Just because *prior trends* are equal doesn't mean that *parallel trends* holds.
- *Parallel trends* is about what the before-after change *would have been* - we can't see that!
- For example, let's say we want to see the effect of online teaching on student test scores, using COVID school shutdowns to get a Before/After
- As of March/April 2020, some schools had gone online (Treated) and others hadn't (Untreated)
- Test score trends were probably pretty similar in the Before periods (Jan/Feb 2020), so prior trends are likely the same
- But LOTS of stuff changed between Jan/Feb and Mar/Apr, like, uh, Coronavirus, lockdowns, etc. not just online teaching! SO *parallel trends* likely wouldn't hold

Concept Checks

- Go back to the Seattle minimum wage effect example from the first Concept Check slide. Clearly state what the parallel trends assumption means in this context.
- It's possible (although perhaps unlikely) that parallel trends can hold even if it looks like the treatment and control groups were trending apart before treatment went into effect. How is this possible?

Example: Pandemic Unemployment Assistance

Pandemic Unemployment Assistance

- During the pandemic the federal government expanded unemployment insurance with the CARES Act set to expire in September 2021
- Starting in May 2021, governors started announcing that they would end the expanded benefits early
- This created a natural experiment where some states ended the benefits early and some did not
- We can use this to estimate the effect of the benefits on employment
 - My co-authors and I did this in [Coombs \(2021\)](#)
- Sadly, we used restricted data, but I will walk you through how to do it
 - You can do your own version with public data! (I almost made this a problem set)

Data

- The data come from Earnin, a fintech company that provides early access to wages
- Have data on bank transactions, earnings, bank account balances, location, gender, and more
- Users are predominantly low-income and liquidity-constrained
 - Critical group to study to learn about effects of unemployment insurance
 - **Big Data point:** Big Data sometimes means you get a non-random sample, but it could be of critical interest for your study
- There are over 3M users, but we sample down

Defining Treatment and Control

- Announcements started in May 2021
- Expirations started in June 2021
- **Treatment:** Live in states that end benefits early
- **Control:** Live in state that does end benefits early
 - Multiple groups (each state)
 - Multiple time periods (each month)
- We define our analysis sample as people who were unemployed and insured at the end of April 2021
 - Why April instead of May?
 - We want to account for people responding to the announcement by going back to work before the expiration
- **Key point:** Expectations can muddy difference-in-difference, so keep an eye out for them

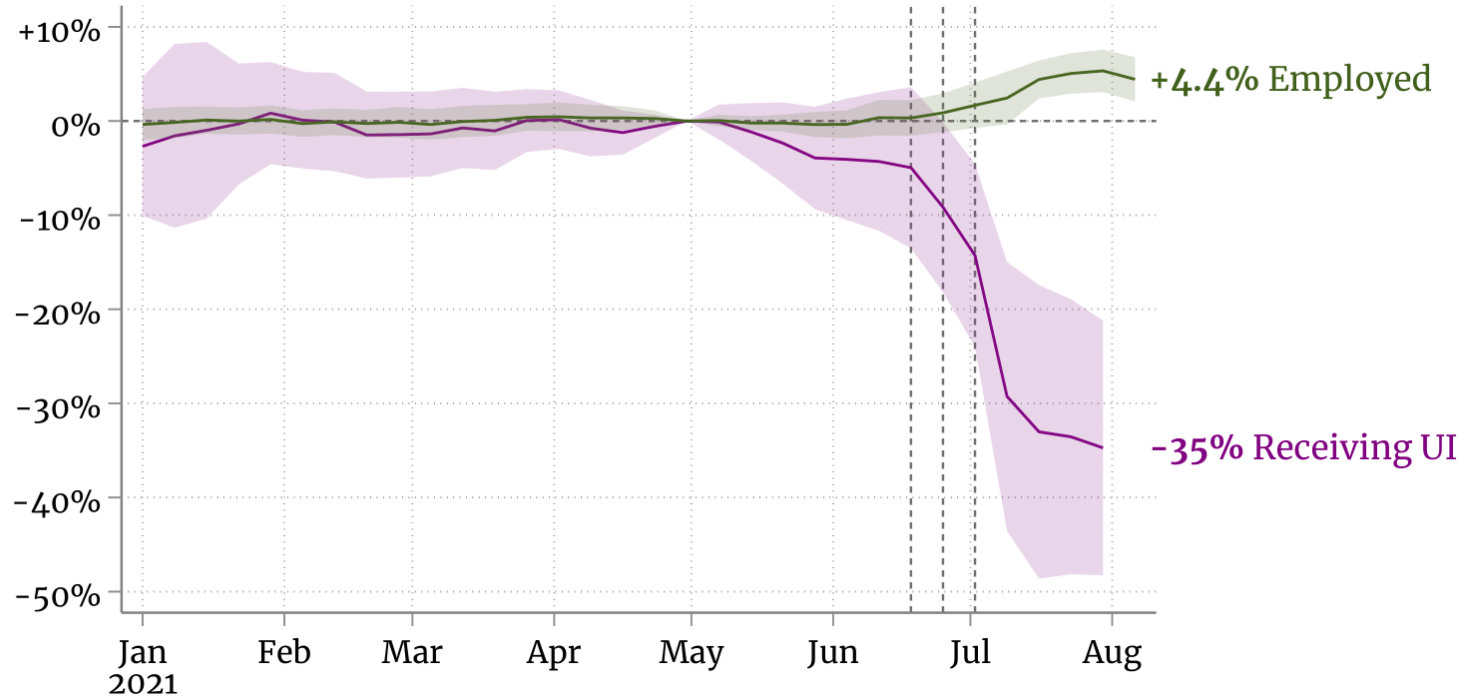
Regression model

$$y_{it} = \alpha_i + \text{week}_t + \sum_{t \neq \text{April 30}} \beta_t \text{Week}_t \times \text{Withdrawal State} + \varepsilon_{it}$$

- Interact each week with the treatment indicator
- Then write that out using the summation term to be more parsimonious
- Wait, why did you omit April 30?
 - Due to multicollinearity, we can't estimate all the week effects and the interaction terms at the same time
 - Have to omit one, practice is to omit the period of or just before the treatment
 - (This gets more complicated if you have multiple treatments at different times)

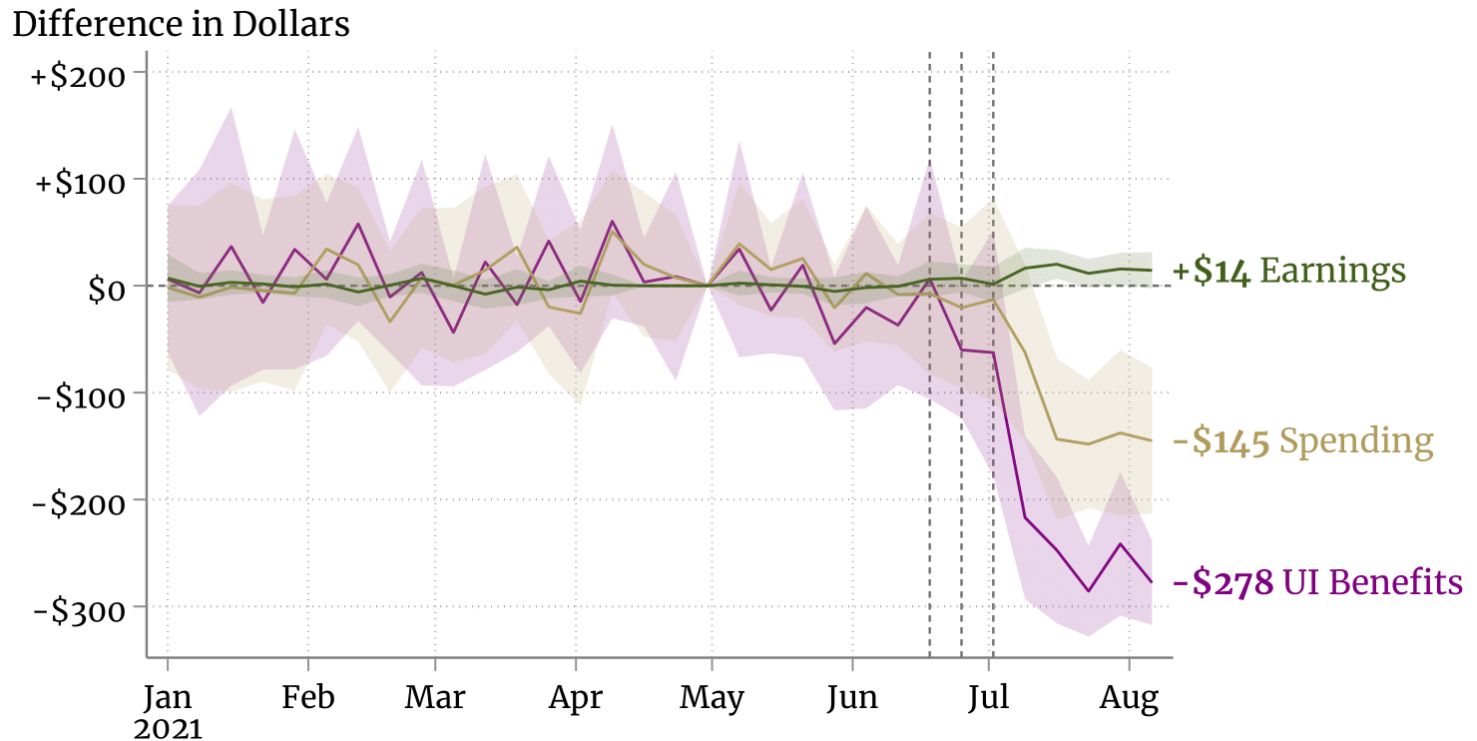
Insured rate vs. employment

Difference in Percentage of Workers



Coombs (2021).

UI inflows, spending, earnings



Coombs (2021).

Other studies

- Other studies have looked at this question using Current Population Survey data (e.g. [Ganong et al. \(2022\)](#), [Holzer et al. \(2023\)](#),)
- They find larger employment effects
 - Reasons: different samples
 - We have a younger, poorer sample
 - Job rationing effects may keep this group out of the labor force longer
 - Younger people may have looked to switch careers
 - Younger people may have younger children
- All that to say, no one study has all the answers!

Before we finish, a warning!

- DID is so nice and simple that it feels like you can get real flexible with it
- But the stuff we're covering in this class - up to TWFE, *relies very strongly* on the assumptions we made. If you break them, the *research design* may hold up, *but the estimator really doesn't* and you need a different estimator

And a Warning!

- TWFE *does not work* if different groups get treated at different times (try: Callaway & Sant'anna (R **did**) or Wooldridge (R **etwfe**))
- If you need to control for stuff to support parallel trends, *just tossing controls in a TWFE doesn't work* (try **did** again, or Sant'anna and Zhao **DRDID**)
- If the outcome is binary, you can't just run TWFE but with logit/probit (just do LPM)
- If the effect gets stronger or weaker over time, the TWFE effect will be weird (do "dynamic DID", see the textbook)
- For now, stick to DID designs where all the Treated groups are treated in the same period

What next?

- Now I want you to go try to implement fixed effects and difference-in-differences models in R
- Navigate to the lecture activity [11a-panel-twfe](#)

Next lecture: Regression Discontinuity Design
