Political Analysis of Social Media Data

Event Studies

Instructor: Gregory Eady
Office: 18.2.10
Office hours: Fridays 13-15

Today

- Event studies
- Video lectures & exercises

Methods for causal inference in the social sciences

- Experiments
- Instrumental variables
- Regression discontinuity designs
- **Difference-in-differences**

Difference-in-differences with social media data

- Allows for examination of the effects of an event or policy change
- Typically with users who are treated as compared to control users who are not
- Or can examine the differential effects of an event on multiple groups of users

Treatment and control case

Can Exposure to Celebrities Reduce Prejudice? The Effect of Mohamed Salah on Islamophobic Behaviors and Attitudes

ALA' ALRABABA'H *Stanford University*
WILLIAM MARBLE *Stanford University*
SALMA MOUSA *Yale University*
ALEXANDRA A. SIEGEL *University of Colorado Boulder*

Can exposure to celebrities from stigmatized groups reduce prejudice? To address this question, we study the case of Mohamed Salah, a visibly Muslim, elite soccer player. Using data on hate crime reports throughout England and 15 million tweets from British soccer fans, we find that after Salah joined Liverpool F.C., hate crimes in the Liverpool area dropped by 16% compared with a synthetic control, and Liverpool F.C. fans halved their rates of posting anti-Muslim tweets relative to fans of other top-flight clubs. An original survey experiment suggests that the salience of Salah's Muslim identity enabled positive feelings toward Salah to generalize to Muslims more broadly. Our findings provide support for the parasocial contact hypothesis—indicating that positive exposure to out-group celebrities can spark real-world behavioral changes in prejudice.

Differential treatment effects

The Profession

The Pandemic and Gender Inequality in Academia

Eunji Kim, *Vanderbilt University, USA*
Shawn Patterson Jr., *Southern Oregon University, USA*

ABSTRACT Has the pandemic exacerbated gender inequality in academia? We provide real-time evidence by analyzing 1.8 million tweets from approximately 3,000 political scientists, leveraging their use of social media for career advancement. Using automated text analysis and difference-in-differences estimation, we find that although faculty members of both genders were affected by the pandemic, the shift to remote work caused women to tweet less often than their male colleagues about professional accomplishments. We argue that these effects are driven by the increased familial obligations placed on women, as demonstrated by the increase in family-related tweets and the more pronounced effects among junior academics. Our evidence demonstrating the gendered shift in professional visibility during the pandemic provides the opportunity for proactive efforts to address disparities that otherwise may take years to manifest.

Difference-in-differences setup

- Must have "panel data": data on which units (individuals, regions, countries) are observed over time
- There is a "shock" to some units (a treatment group) at a specific period in time, but not others (a control group)
  - e.g. A new policy is implemented
  - e.g. An event occurs in some place, but not others
- Assume that treatment units would have followed the same trend as the control group were it not for the shock (the counter-factual)
  - Parallel trends assumption

Card and Krueger (1994): The classic DD setup

- Debate among economists about whether increasing the minimum wage causes an increase in unemployment
- At the time, there is cross-sectional evidence that this is true
- But US states do not select a minimum wage at random, so cross-sectional designs not appropriate

Card and Krueger's (1994): The solution

- Compare a treatment and control case over time
- New Jersey raised its minimum wage in April 1992
- Card and Krueger (1994) compare employment in New Jersey's fast food industry to that of neighboring Pennsylvania before and after the minimum wage increase
- Result: If anything, a *positive* effect on employment

Card and Krueger's (1994): Two-period diff-in-diff

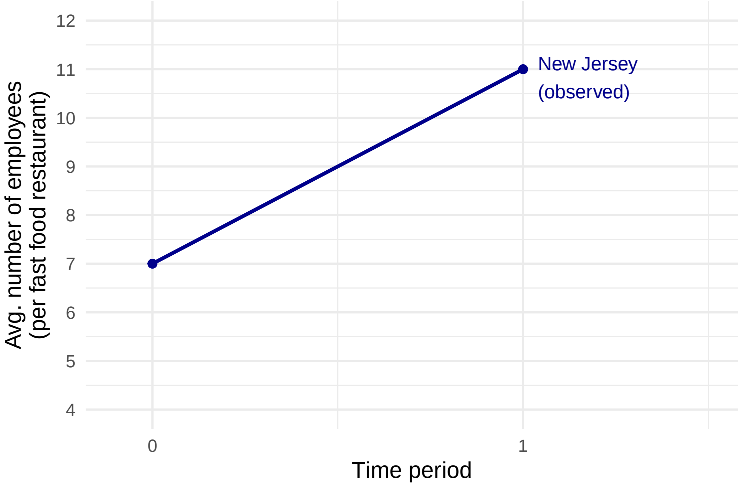
- Two units:
  - Treatment case: New Jersey
  - Control case: Pennsylvania
- Four observations
  - Pre-treatment ($t = 0$)
    - Minimum wage in New Jersey (untreated)
    - Minimum wage in Pennsylvania (untreated)
  - Post-treatment ($t = 1$)
    - Minimum wage in New Jersey (treated)
    - Minimum wage in Pennsylvania (untreated)

Difference-in-differences: Compare the difference in employment between New Jersey and Pennsylvania at $t = 0$ to the difference at $t = 1$

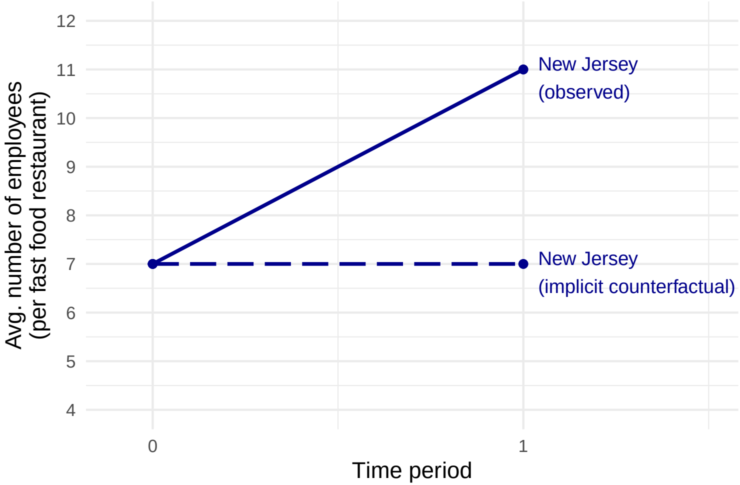
Why are we comparing the difference in employment between two states?

- We need a way to create a counter-factual comparison for New Jersey
- We will assume that—if no new policy were implemented—changes in the number of employees for New Jersey and Pennsylvania would move *in parallel*

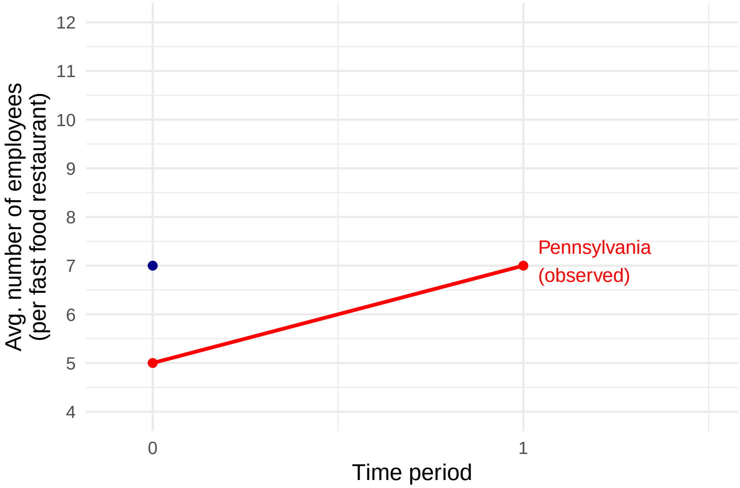
Naive pre-post comparison $(11 - 7 = 4?)$



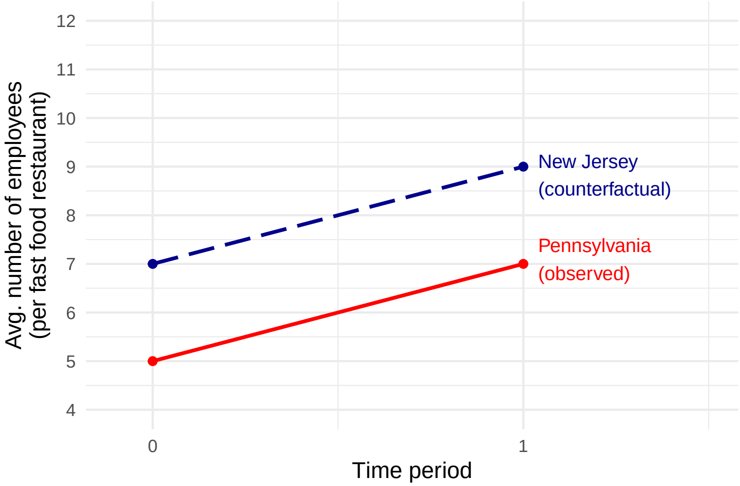
Naive pre-post comparison ($11 - 7 = 4$?)



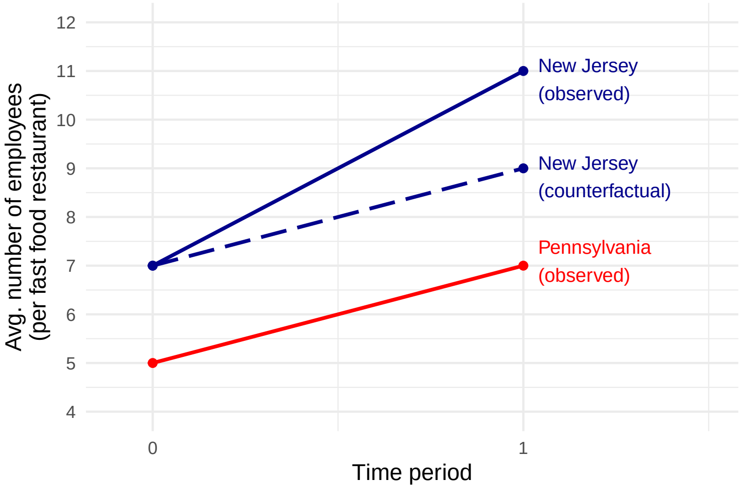
Pennsylvania as a control



Pennsylvania versus "New Jersey"



Pennsylvania versus New Jersey



Difference-in-differences estimate

- New Jersey (treated):
  - Before: 7
  - After: 11
  - $Difference_{NJ}$: $11 - 7 = 4$
- Pennsylvania (control):
  - Before: 5
  - After: 7
  - $Difference_{PA}$: $7 - 5 = 2$
- The difference in these two differences? $4 - 2 = 2$
- This is the *difference-in-differences* estimate of the effect of a minimum wage increase on employment

A bit more formally

Difference for New Jersey from $t = 0$ to $t = 1$:

$$\left(\bar{Y}_{\mathrm{NJ,t}\,=\,1} - \bar{Y}_{\mathrm{NJ,t}\,=\,0}\right) \tag{1}$$

Difference for Pennsylvania from $t = 0$ to $t = 1$:

$$\left(\bar{Y}_{\mathrm{PA,t}\,=\,1} - \bar{Y}_{\mathrm{PA,t}\,=\,0}\right) \tag{2}$$

The difference-in-differences estimate is:

$$\left(\bar{Y}_{\mathrm{NJ,t}\,=\,1} - \bar{Y}_{\mathrm{NJ,t}\,=\,0}\right) - \left(\bar{Y}_{\mathrm{PA,t}\,=\,1} - \bar{Y}_{\mathrm{PA,t}\,=\,0}\right) \tag{3}$$

The big assumption that allows us to use an untreated unit as a control is the "parallel trends assumption"

That, had there been no change in policy:

$$\overbrace{(Y(0)_{post}|T=1)-Y(0)_{pre}|T=1)}^{\substack{\text{Difference in New Jersey employment}\\ \text{if } no \text{ minimum wage policy change}}} -$$

$$\underbrace{(Y(0)_{post}|T=0)-Y(0)_{pre}|T=0)}_{\substack{\text{Difference in Pennsylvania employment}\\ \text{if } no \text{ minimum wage policy change}}}$$

We need to assume that this is equal to zero (else the diff-in-diff estimate is biased)

Why might the parallel trends assumption be broken?

- Something else happens at the same time as the treatment that affects the groups differently
  - e.g. A big McDonald's ad campaign in New Jersey
- Other shocks or events
  - Macro- or micro-level economic forces affect Pennsylvania differently from New Jersey
- Longer term trends are different in general

Choose of control unit(s) thus matters

- Card and Krueger (1994) recognized this:
  - New Jersey and Pennsylvania have similar economic composition
  - Same weather
  - Same region, so similar economic or other shocks
- Nevertheless this might give (undue) discretion to researchers
  - Can also automate the selection of control comparison cases with “synthetic control” methods (Adadie et al. 2003, 2010, 2015)

This is the canonical two-period diff-in-diff. But...

- Often we have many periods, so can't run a simple regression for just two cases
- Often treatment timing varies (e.g. a minimum wage increase is implemented in different states at different times)
- We thus need a *generalized* difference-in-differences model
- So let's look at one example as an application...

COVID-19 and gender inequality in academia

The Profession

The Pandemic and Gender Inequality in Academia

Eunji Kim, *Vanderbilt University, USA*
Shawn Patterson Jr., *Southern Oregon University, USA*

ABSTRACT Has the pandemic exacerbated gender inequality in academia? We provide real-time evidence by analyzing 1.8 million tweets from approximately 3,000 political scientists, leveraging their use of social media for career advancement. Using automated text analysis and difference-in-differences estimation, we find that although faculty members of both genders were affected by the pandemic, the shift to remote work caused women to tweet less often than their male colleagues about professional accomplishments. We argue that these effects are driven by the increased familial obligations placed on women, as demonstrated by the increase in family-related tweets and the more pronounced effects among junior academics. Our evidence demonstrating the gendered shift in professional visibility during the pandemic provides the opportunity for proactive efforts to address disparities that otherwise may take years to manifest.

What were the effects of the COVID-19 pandemic on family obligations among women and men in academia?

- Collect users names of political scientists who follow at least 1 of 5 major political science accounts ($n = 2,912$)
- Manually search website pages of each to determine gender, rank, and institutional affiliation
- Collect all tweets from each person from June 1, 2019 onward

Measuring work- and family-related topics

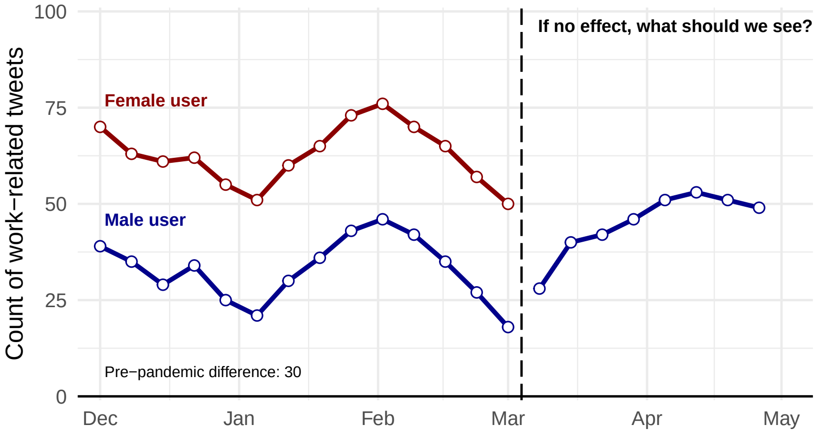
- Keyword search for family and work terms
- Hand coded 100 of the tweets with these keyword and expanded the keyword list based on these
- For each user's Twitter feed, calculate the number of family-related and work-related tweets
  - i.e. each row is the count (and proportion) of family-related tweets in a given week by a given user

Example data

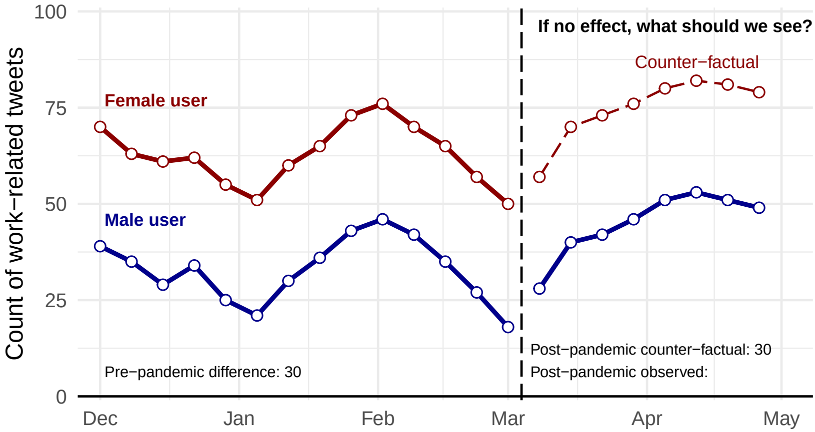
week	user_id	female	rank	num_family	prop_family	num_work	prop_work	pandemic
2019-06-01	1	1	Full prof.	30	0.15	40	0.2	0
2019-06-08	1	1	Full prof.	20	0.1	10	0.05	0
2019-06-15	1	1	Full prof.	80	0.4	10	0.05	0
2019-06-22	1	1	Full prof.	25	0.12	30	0.15	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2020-05-02	2912	1	Asst prof.	80	0.4	10	0.05	1
2020-05-09	2912	1	Asst prof.	25	0.12	30	0.15	1

- i.e. each row is the number/proportion of tweets for one specific user in a given week

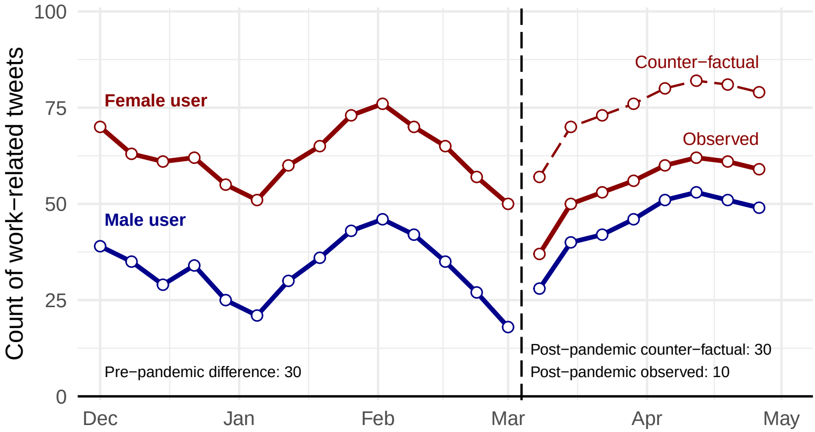
What is being estimated in a diff-in-diff framework?



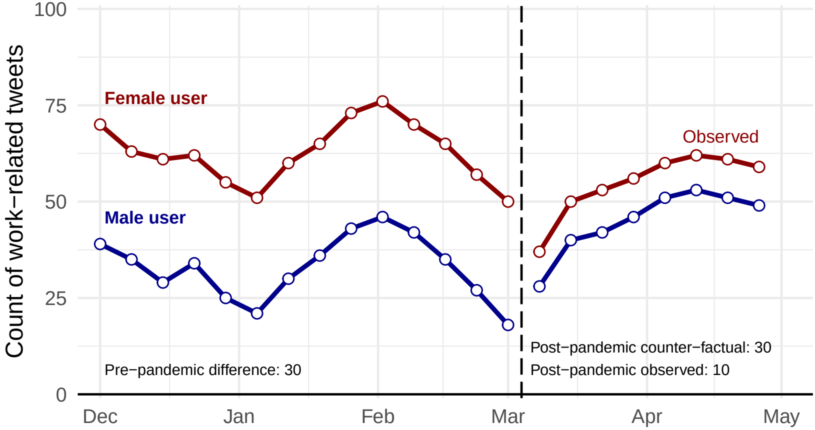
What is being estimated in a diff-in-diff framework?



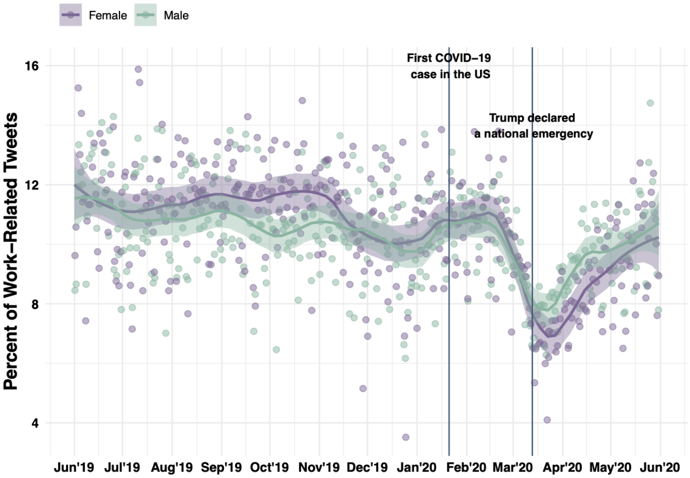
What is being estimated in a diff-in-diff framework?



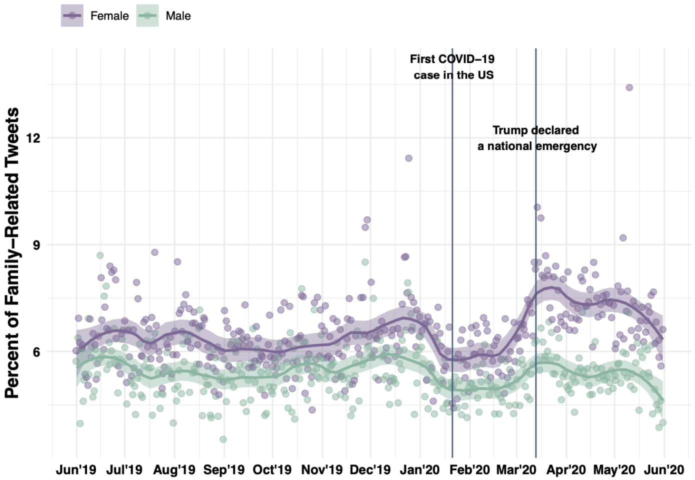
What is being estimated in a diff-in-diff framework?



Work-related tweets for women and men



Family-related tweets for women and men



Okay, so how do we estimate a diff-in-diff across all users?

$$y_{it} = \delta_i + \mu_t + \beta(\text{pandemic}_t \times \text{woman}_i) + \epsilon_{it}$$

Okay, so how do we estimate a diff-in-diff across all users?

$$y_{it} = \delta_i + \mu_t + \beta(\text{pandemic}_t \times \text{woman}_i) + \epsilon_{it}$$

- y_{it}: outcome (family-related tweets/work-related tweets)

Okay, so how do we estimate a diff-in-diff across all users?

$$y_{it} = \delta_i + \mu_t + \beta(\text{pandemic}_t \times \text{woman}_i) + \epsilon_{it}$$

- y_{it}: outcome (family-related tweets/work-related tweets)
- δ_i: user fixed effect

Okay, so how do we estimate a diff-in-diff across all users?

$$y_{it} = \delta_i + \mu_t + \beta(\text{pandemic}_t \times \text{woman}_i) + \epsilon_{it}$$

- y_{it}: outcome (family-related tweets/work-related tweets)
- δ_i: user fixed effect
- μ_t: week fixed effect

Okay, so how do we estimate a diff-in-diff across all users?

$$y_{it} = \delta_i + \mu_t + \beta(\text{pandemic}_t \times \text{woman}_i) + \epsilon_{it}$$

- y_{it}: outcome (family-related tweets/work-related tweets)
- δ_i: user fixed effect
- μ_t: week fixed effect
- β: effect of the pandemic on women relative to men

Okay, so how do we estimate a diff-in-diff across all users?

$$y_{it} = \delta_i + \mu_t + \beta(\text{pandemic}_t \times \text{woman}_i) + \epsilon_{it}$$

- y_{it}: outcome (family-related tweets/work-related tweets)
- δ_i: user fixed effect
- μ_t: week fixed effect
- β: effect of the pandemic on women relative to men
- Note that Kim and Patterson Jr. have an α in their model specification, but that will just drop out
- Cluster your standard errors at the unit level

Regression results

Table 3: The Pandemic Effect on Family- and Work-Related Tweets

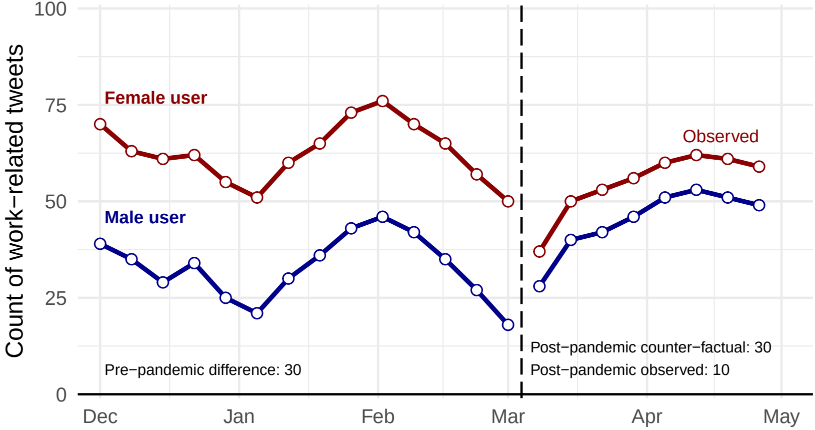
	Family	Work	Family	Work	Family	Work	Family	Work
	All Faculty		Assistant		Associate		Full	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female*Pandemic	0.967***	-1.354***	1.268***	-1.631**	0.811*	-1.188*	0.573	-0.895
	(0.220)	(0.324)	(0.353)	(0.498)	(0.387)	(0.579)	(0.406)	(0.630)
Individual Fixed Effect?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effect?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	100,152	100,152	43,052	43,052	31,735	31,735	25,365	25,365
R^2	0.114	0.181	0.107	0.154	0.116	0.200	0.128	0.209

Note: Standard errors are in parentheses. * p<0.05; ** p<0.01; *** p<0.001

Validity depends heavily on the parallel trends assumption, however

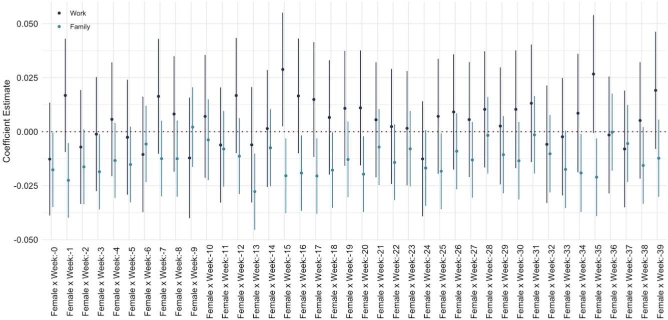
- Our estimate is valid if our treatment group would have changed in the same way as our control group were it not for the event or policy change
- This is fundamentally unknowable...
- But we can check this indirectly by examining whether the trends between groups is parallel *before* the event or policy change
- i.e. in each time period, does the difference between groups more or less stay the same over time?

Look at the pre-trend differences:



Check for pre-pandemic parallel trends by examining differences week by week before the intervention

Figure 3: **Pre-Treatment Gender Difference in Work- and Family-Related Tweets (%)**



Note: This figure reports the estimated coefficients of Female interacted with the dummies for each week pre-treatment. The benchmark week is Week 40 (i.e the first treated week). Individual and time fixed effects are included. Vertical lines represent the 95% CI and the dots indicate the estimated coefficient of the interaction term. Navy indicates the results for the work-related tweets, and lighter blue indicates the results for the family-related tweets. Null results suggest there was no consistent pre-treatment difference in trends.

Maybe something happens in March in general to cause this effect?

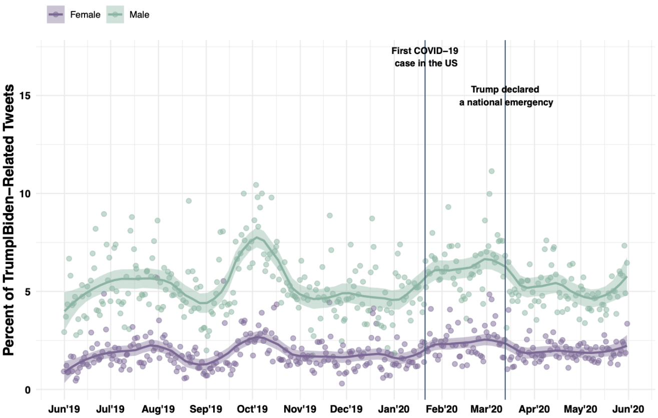
Table 4: The Absence of the Seasonality Effect in Gender Difference

	2019		2020	
	Family	Work	Family	Work
	(1)	(2)	(3)	(4)
Female*Pandemic	0.014	-0.497	1.108***	-1.231***
	(0.277)	(0.431)	(0.263)	(0.372)
Individual Fixed Effect?	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
Observations	40,518	40,518	46,257	46,257
R^2	0.159	0.222	0.171	0.218

Note: Standard errors are in parentheses. * p<0.05; ** p<0.01; *** p<0.001

What if we ran the model on an outcome where we shouldn't expect to find any effect?

Figure 4: Daily Trends in % Trump- and Biden-Related Tweets



What if we ran the model on an outcome where we shouldn't expect to find any effect?

Table 5: The Lockdown Effect on Trump- and Biden-Related Tweets

	All Faculty	Assistant	Associate	Full
	(1)	(2)	(3)	(4)
Female*Pandemic	0.087	0.003	0.104	0.259
	(0.146)	(0.173)	(0.258)	(0.370)
Individual Fixed Effect?	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
Observations	101,408	43,589	32,135	25,684
R^2	0.254	0.183	0.215	0.320

Note: Standard errors are in parentheses. * p<0.05; ** p<0.01; *** p<0.001

Event study models

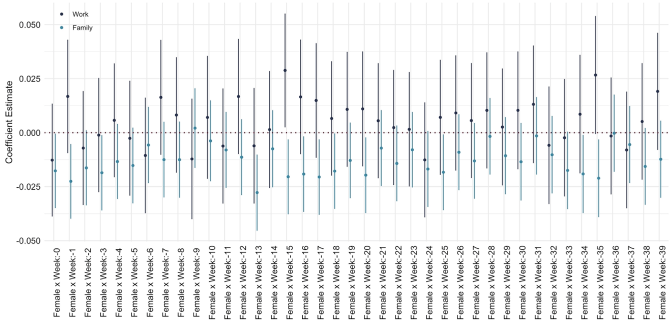
- Diff-in-diff models calculate an effect as a weighted average of treated and control units pre- and post-event/policy
- But what if we want to see the dynamics of an effect?
- How long does it last?
- Does it occur immediately?

Event study models

- Event study models are effectively just difference-in-differences per time period
- In Kim & Patterson Jr., compare each week after the lock-downs relative to a lock-down baseline ...

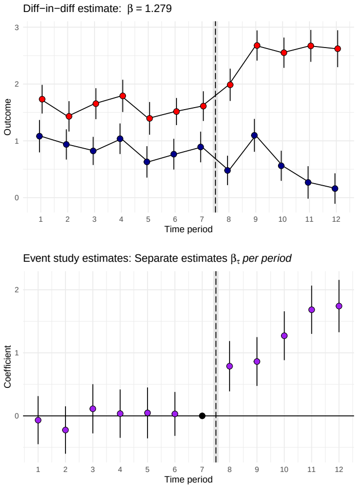
For example, recall...

Figure 3: **Pre-Treatment Gender Difference in Work- and Family-Related Tweets (%)**



Note: This figure reports the estimated coefficients of Female interacted with the dummies for each week pre-treatment. The benchmark week is Week 40 (i.e the first treated week). Individual and time fixed effects are included. Vertical lines represent the 95% CI and the dots indicate the estimated coefficient of the interaction term. Navy indicates the results for the work-related tweets, and lighter blue indicates the results for the family-related tweets. Null results suggest there was no consistent pre-treatment difference in trends.

Event studies calculate differences between treatment and control per time period



Event study application

Do Violent Protests Affect Expressions of Party Identity? Evidence from the Capitol Insurrection

GREGORY EADY *University of Copenhagen, Denmark*
FREDERIK HJORTH *University of Copenhagen, Denmark*
PETER THISTED DINESEN *University College London, United Kingdom, and University of Copenhagen, Denmark*

The insurrection at the United States Capitol on January 6, 2021, was the most dramatic contemporary manifestation of deep political polarization in the United States. Recent research shows that violent protests shape political behavior and attachments, but several questions remain unanswered. Using day-level panel data from a large sample of US social media users to track changes in the identities expressed in their Twitter biographies, we show that the Capitol insurrection caused a large-scale decrease in outward expressions of identification with the Republican Party and Donald Trump, with no indication of reidentification in the weeks that followed. This finding suggests that there are limits to party loyalty: a violent attack on democratic institutions sets boundaries on partisanship, even among avowed partisans. Furthermore, the finding that political violence can deflect copartisans carries the potential positive democratic implication that those who encourage or associate themselves with such violence pay a political cost.

What are the consequences of violent protests on political behavior and attachments? Two recent studies:

- Proximity to black-led protests caused increased support for restrictive policies and support for the Republican Party (Wasow 2020)
- Proximity to LA riots led to liberal shift in policy support, and increase in support for Democratic Party (Enos et al. 2019)

Unanswered questions

- Politics are now heavily nationalized, so what are broader effects of violent protest?
- What are the consequences of violent protest when conducted by those on the political right?
- Do these effects occur quickly, or only after longer sustained elite politicization?

Research setup

- Examine the behavioral reaction to the Capitol insurrection by studying online de-identification with the Republican Party (and Donald Trump)
- A hard test of the scope conditions of the ‘unmovable’ character of (expressed) partisanship in the US

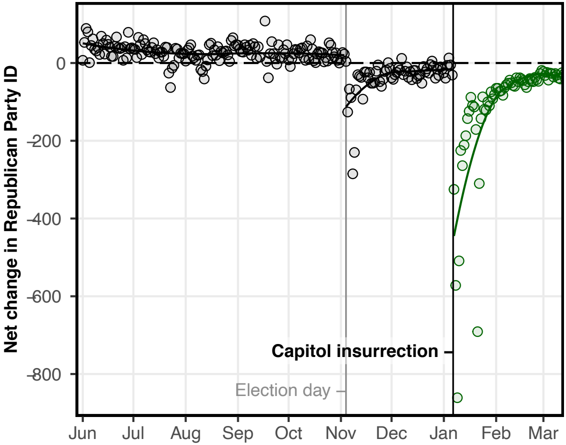
Research setup

- Day-level panel data of the profiles of 3.4 million active US Twitter users (June 1, 2020+)
- Follow 1+ major US news organization (from MSNBC to Brietbart)
- ∼1 billion user-day observations of Twitter bios (profiles)
- Keyword expansion to identify explicitly partisan terms in users' bios

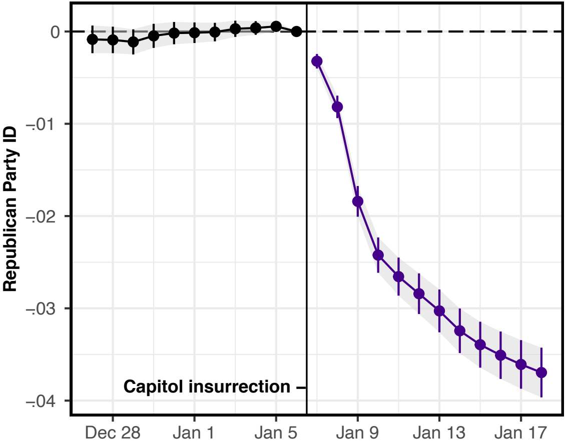
A stylized example

date	profile text
2020-06-01	Proud Texan Republican! Grandmother, mother, Christian, #MAGA
2020-06-02	Proud Texan Republican! Grandmother, mother, Christian, #MAGA
2020-06-03	Proud Texan Republican! Grandmother, mother, Christian, #MAGA
2020-06-04	Proud Texan! Grandmother, Mother, Christian, and proud American
2020-06-05	Proud Texan! Grandmother, Mother, Christian, and proud American
⋮	⋮

Net change in Republican ID over time



Event study



* Not driven by *increase* in Democratic identification (is flat)

Okay, so how do we estimate this event study model?

$$y_{it} = \alpha_i + \lambda_t + \sum_{t \neq 0} \beta_t \text{Republican}_i + \epsilon_{it},$$

Okay, so how do we estimate this event study model?

$$y_{it} = \alpha_i + \lambda_t + \sum_{t \neq 0} \beta_t \text{Republican}_i + \epsilon_{it},$$

- y_{it}: outcome (partisan terms in Twitter bio)

Okay, so how do we estimate this event study model?

$$y_{it} = \alpha_i + \lambda_t + \sum_{t \neq 0} \beta_t \text{Republican}_i + \epsilon_{it},$$

- y_{it}: outcome (partisan terms in Twitter bio)
- α_i: user fixed effect

Okay, so how do we estimate this event study model?

$$y_{it} = \alpha_i + \lambda_t + \sum_{t \neq 0} \beta_t \text{Republican}_i + \epsilon_{it},$$

- y_{it}: outcome (partisan terms in Twitter bio)
- α_i: user fixed effect
- λ_t: day fixed effect

Okay, so how do we estimate this event study model?

$$y_{it} = \alpha_i + \lambda_t + \sum_{t \neq 0} \beta_t \text{Republican}_i + \epsilon_{it},$$

- y_{it}: outcome (partisan terms in Twitter bio)
- α_i: user fixed effect
- λ_t: day fixed effect
- β_t: effect of the insurrection on Republicans including partisan terms (relative to Democrats)

Okay, so how do we estimate this event study model?

$$y_{it} = \alpha_i + \lambda_t + \sum_{t \neq 0} \beta_t \text{Republican}_i + \epsilon_{it},$$

- y_{it}: outcome (partisan terms in Twitter bio)
- α_i: user fixed effect
- λ_t: day fixed effect
- β_t: effect of the insurrection on Republicans including partisan terms (relative to Democrats)
- Note in the sum there is no β for $t = 0$ because $t = 0$ (the day right before the insurrection) is our baseline

Although this isn't in the paper, one could fit a standard diff-in-diff model

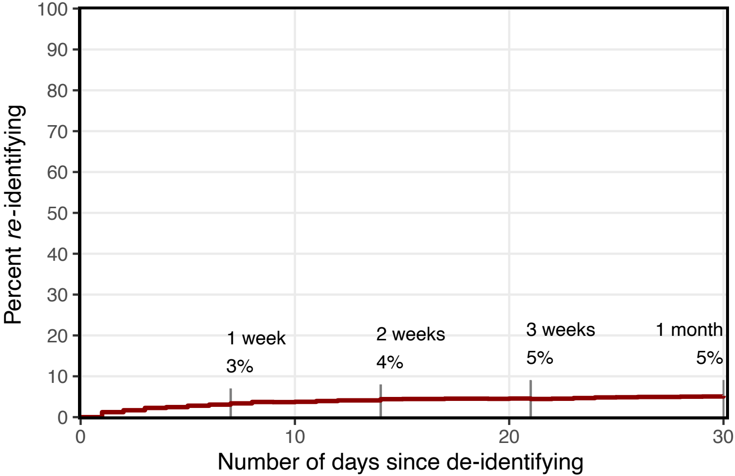
$$y_{it} = \alpha_i + \lambda_t + \beta(\text{Insurrection}_t \times \text{Republican}_i) + \epsilon_{it},$$

- y_{it}: outcome (partisan terms in Twitter bio)
- α_i: user fixed effect
- λ_t: day fixed effect
- β: effect of the insurrection on Republicans including partisan terms (relative to Democrats)
- Note that this compares the difference pre-insurrection to the post-insurrection period *overall* rather than per day

Robustness

- Fear of prosecution?
  - Remove any user who deleted/scrubbed any tweets on day of de-identification.
- Because Twitter deleted QAnon in weeks afterward?
  - Remove any user who was deleted during time period of interest.
  - Also, as above, any user who scrubbed timeline.
- Just Trump-related?
  - Same result (↓ magnitude) if use party-only terms
- But isn't the effect just temporary? ...

Effect duration (re-identifiers)



Conclusions

- Within 3 weeks after the insurrection, 1 in 14 previously Republican-identifying users had removed partisan terms
- Democracy-threatening violence can set boundaries on partisanship, even among avowed partisans
- Positive democratic implication that those who encourage political violence may pay a political cost by way of partisan de-mobilization

Conclusions

- Often good to apply both types of models in a paper
- The parallel trends assumption is extremely important
- Fortunately, event study models allow you to visually check pre-treatment parallel trends
- If the trends are *not* parallel, look into modeling unit-level trends (not difficult to do)
  - Ask me about it if you ever want to do this
- There is a massive literature on diff-in-diff and event study models
  - E.g. synthetic control (like in Alrababa'h et al.)
  - E.g. staggered treatment (units treated in different periods)