Advanced Quantitative Methods

Event Studies & Synthetic Control

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

Today

- Preview of time-varying treatments in diff-in-diff
- Event studies
- Synthetic control

Standard difference-in-differences

- Classic case is when some units become treated in a given period and others are not treated
- $y_{it} = \alpha_i + \gamma_t + \beta \text{Treatment}_{it} + \epsilon_{it}$
  - β is our estimate of the difference-in-differences comparing the pre- to the post-treatment period
- But what if treatment assignment varies over time?
  - e.g. a policy is implemented in one state in 2019, and then in two other states in 2020, and so forth

e.g. Classic case (treatment at the same time):

id	year	treatment	outcome
1	2000	0	34
1	2001	0	25
1	2002	1	27
1	2003	1	30
1	2004	1	24
2	2000	0	78
2	2001	0	68
2	2002	1	68
2	2003	1	71
2	2004	1	89
3	2000	0	20
3	2001	0	13
3	2002	0	9
3	2003	0	30
3	2004	0	26

e.g. Time-varying treatment

id	year	treatment	outcome
1	2000	0	34
1	2001	1	25
1	2002	1	27
1	2003	1	30
1	2004	1	24
2	2000	0	78
2	2001	0	68
2	2002	0	68
2	2003	1	71
2	2004	1	89
3	2000	0	78
3	2001	0	68
3	2002	0	68
3	2003	0	71
3	2004	0	89

We can still use our standard generalized diff-in-diff with time-varying treatments

$$y_{it} = \alpha_i + \gamma_t + \beta \text{Treatment}_{it} + \epsilon_{it}$$

- There are complications, however...
- If you have a time-varying treatment, there are additional robustness checks and models to use:
  - https://andrewcbaker.netlify.app/2019/09/25/difference-in-differences-methodology/
  - https://blogs.worldbank.org/impactevaluations/what-are-we-estimating-when-we-estimate-difference-differences
- But next week we will go over this in more depth

Event studies

- Standard diff-in-diff designs estimate an overall effect
- But we can fit a diff-in-diff model more flexibly with an "event study" model
- Event studies allow one to see the *dynamics* of an effect, i.e. how its magnitude changes over time
- They also allow us to examine pre-trends by providing a visual check on the parallel trends assumption
- It works by constructing a diff-in-diff estimate for every period in your data using "leads" and "lags"

If treatment is assigned in only one period:

id	year	treatment	outcome
1	2000	0	34
1	2001	0	25
1	2002	1	27
1	2003	1	30
1	2004	1	24
2	2000	0	78
2	2001	0	68
2	2002	1	68
2	2003	1	71
2	2004	1	89
3	2000	0	20
3	2001	0	13
3	2002	0	9
3	2003	0	30
3	2004	0	26

If treatment assignment varies over time, we need to construct a new variable

id	year	treatment	outcome	time_to_treatment (τ)
1	2000	0	34	0
1	2001	1	25	1
1	2002	1	27	2
1	2003	1	30	3
1	2004	1	24	4
2	2000	0	78	-2
2	2001	0	68	-1
2	2002	0	68	0
2	2003	1	71	1
2	2004	1	89	2
3	2000	0	78	0
3	2001	0	68	0
3	2002	0	68	0
3	2003	0	71	0
3	2004	0	89	0

How an event study works

- It constructs a difference-in-difference *in every time period*
- What is the difference between treatment and control units in time period 1, and in period 2, in period 3, etc.
- Each time period difference is compared to a baseline difference
  - Typically the time period right before the treatment

Some notation

Standard diff-in-diff model:

$$y_{it} = \alpha_i + \gamma_t + \beta D_{it} + \epsilon_{it}$$

Event study model ((a.) and (b.) are the same, but with different notation):

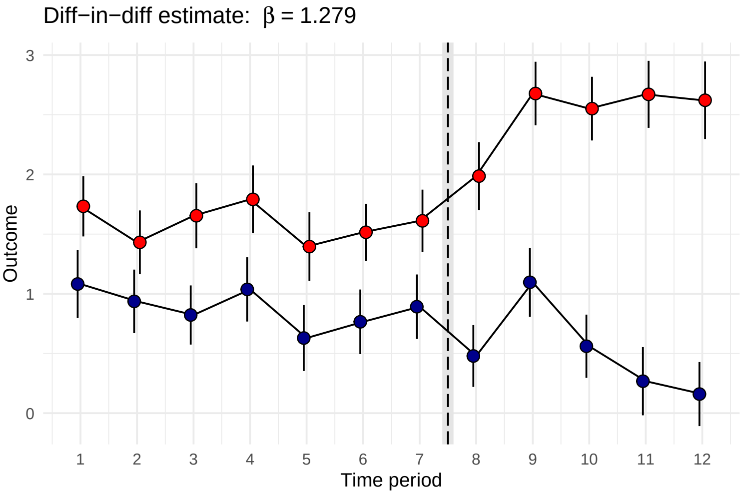
$$y_{it} = \alpha_i + \gamma_t + \sum_{\tau=-q}^{m} \beta_\tau D_{i\tau} + \epsilon_{it}, \tag{a.}$$

$$y_{it} = \alpha_i + \gamma_t + \sum_{\tau=-q}^{-1} \beta_\tau D_{i\tau} + \sum_{\tau=0}^{m} \delta_\tau D_{i\tau} + \epsilon_{it}, \tag{b.}$$

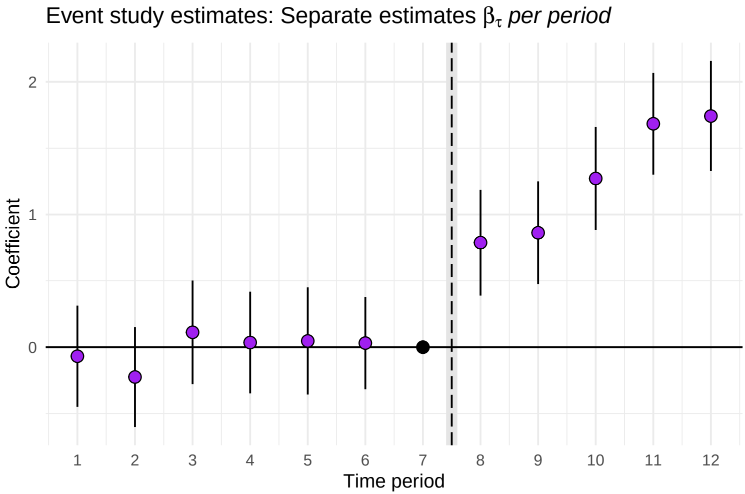
where q is the number of leads (number of time periods before the treatment), and m is the number of lags (number of time periods after the treatment).

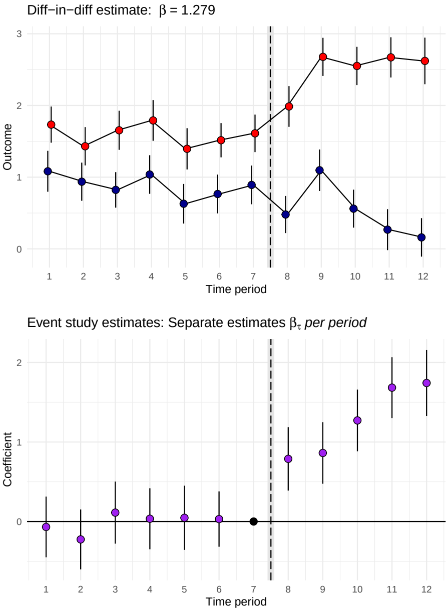
You thus have a *set* of parameters β_τ that are separate diff-in-diff estimates...

Standard diff-in-diff graph & estimate



Event study graph




Diff−in−diff estimate: β = 1.279
Outcome
0
1
2
3
1 2 3 4 5 6 7 8 9 10 11 12
Time period
Event study estimates: Separate estimates βτ per period
Coefficient
0
1
2
1 2 3 4 5 6 7 8 9 10 11 12
Time period

Example write-up of event study in an article:

American Economic Review 2023, 113(6): 1424–1460
https://doi.org/10.1257/aer.20201867

The Birth of a Nation: Media and Racial Hate†

By Desmond Ang*

This paper documents the impact of popular media on racial hate by examining the first American blockbuster: 1915's The Birth of a Nation, *a fictional portrayal of the KKK's founding rife with racist stereotypes. Exploiting the film's five-year "road show," I find a sharp spike in lynchings and race riots coinciding with its arrival in a county. Instrumenting for road show destinations using the location of theaters prior to the movie's release, I show that the film significantly increased local Klan support in the 1920s. Road show counties continue to experience higher rates of hate crimes and hate groups a century later.* (*JEL* J15, K42, L82, N31, N32, N41, Z13)

From Ang (2023):

Nonetheless, it is important to note that the event study design used to examine short-run effects does not require unobserved factors correlated with racism to be uncorrelated with treatment. The model instead relies on a parallel trends assumption and accounts for level differences between treatment and control areas by examining changes in acts of racial violence before and after the film's arrival. In particular, I estimate the following equation on weekly panel data from 1913 to 1922 for all counties in the continental United States:[13]

$$y_{c,t} = \delta_c + \lambda_{s,t} + \sum_{\tau=-6}^{6} \beta_\tau Show_\tau + \epsilon_{c,t}. \tag{1}$$

Here, $y_{c,t}$ is an indicator for whether a lynching (race riot) occurred in county c at week t, δ_c are county fixed effects, and $\lambda_{s,t}$ are state-week fixed effects that flexibly control for state-wide trends or shocks. To reduce noise stemming from the rarity of lynchings, the coefficients of interest (β_τ) are on a vector of dummies corresponding to relative *months* to *The Birth of a Nation*'s arrival in a county.[14] The omitted period is the last month before the movie's arrival in a county and $Show_6$ ($Show_{-6}$) is set to 1 for periods that are 6 or more months after (before) a showing. Standard errors are clustered by state.

Results

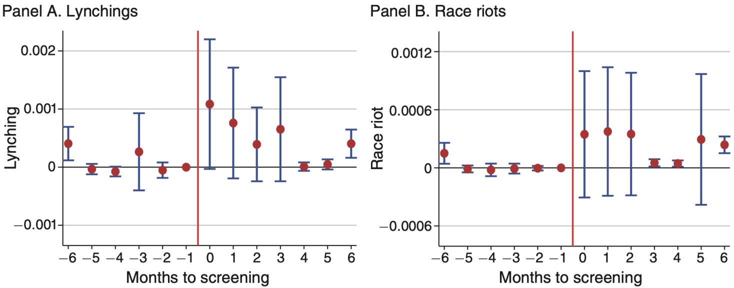


FIGURE 3. CONTEMPORANEOUS EFFECTS ON RACIAL VIOLENCE

Notes: Figure shows event study estimates and 95 percent confidence intervals from equation (1) of the effect of BON screenings on racial violence. Outcomes of interest are indicators for whether a county experienced a lynching (mean = 0.0003) or race riot (mean = 0.00009) in a week. Sample spans 1913 to 1922. Unit of observation is county-week. Months to treat represent five-week intervals centered on the week of BON's arrival in a county. Standard errors clustered by state.

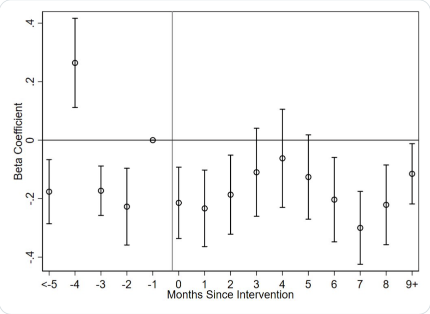
Results write-up

It is important to note that the lynching and riot effects are highly imprecise and are not statistically significant at conventional levels ($p = 0.06$ for $\beta_0^{lynching}$; $p = 0.30$ for β_0^{riot}). This is largely due to the scarcity of lynchings and riots. From 1913 to 1922, a period spanning the Red Summer and the peak of race riots in the United States, there were fewer than 20 race riots and 60 lynchings per year nationwide. As such, the effects are driven by a small number of events: 15 treatment counties experienced a lynching or race riot in the year after the film's arrival. Nevertheless, as I demonstrate in the next subsection, the sharp increase in racial violence is preserved across a host of robustness checks, and randomization inference yields highly significant p-values ($p < 0.01$ for lynchings and $p < 0.05$ for riots).

Research is hard

Analisa Packham
@analisapackham

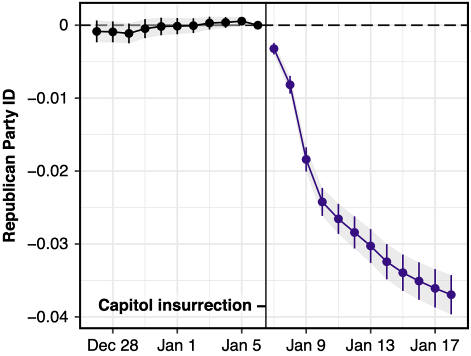
On Twitter, it's easy to think everyone else's research is spectacular and no one else is hitting roadblocks. So today I'm sharing a DiD graph I made that...isn't so spectacular. To younger scholars out there: Research is hard! Things don't always work perfectly. Keep at it.



11:40 AM · Jul 7, 2021 · Twitter Web App

But sometimes everything looks exceptionally clear

FIGURE 2. Event Study Estimates (with 95% CIs)



Note: Data were collected each morning, and thus observations on January 6 (before vertical line) are preinsurrection. Standard errors are clustered at the user level.

Event studies

- Easy to implement in R
- Intuitive to understand
- An essential visual check on pre-intervention parallel trends
- Captures the dynamics of an effect
- Increasingly used in political science
  - These will soon be required in all diff-in-diff studies (like in economics)
- Can be used as the primary model, or in concert with a standard diff-in-diff model

Synthetic control

- How can we examine the effect of an event that affects a single unit?
  - What was the effect of German reunification on the economic performance of West Germany? (Abadie, Diamond & Hainmueller, 2010)
  - What was the effect of the exodus of Cubans from Mariel Harbor to Miami on the local economy? (David Card 1990)
  - What was the effect of terrorism in Basque Country in the 1960s on economic growth? (Abadie & Gardeazabal 2003)
  - What was the effect of the civil war in Congo on deforestation? (Kikuta 2020)

Synthetic control

- Counter-factually, we want to know *what would have happened* to the treated unit had it not been treated
- We cannot observe this, i.e. the "fundamental problem of causal inference"
- Synthetic control methods are designed to construct a counter-factual by using a weighted average of other units to mimic the counter-factual of our unit of interest

Synthetic control

- Like diff-in-diff, one needs panel data for the unit of interest, and other units which we might think of as similar
- Each unit is observed at the same points in time over the time period of interest
- Then, in one period, the unit of interest is treated, and the others remain as controls
  - Comparing West Germany as the treated unit to all other OECD countries from 1960 to 2003
- Need predictors of the outcome of interest as well—*not* just the data for the outcome from each unit

Synthetic control as weighted combination

- The synthetic control is represented as a vector of weights $W = (w_2, w_3, \ldots, w_{J+1})$,
  - where there are $J+1$ total units in the data, and $j = 1$ is the treated unit
  - The value of the weights w are constrained to be greater or equal to 0 and less than or equal to 1
  - And the weights are constrained to sum to 1
- X are a set of pre-intervention characteristics from each unit that we will match on
- $X_1 - X_0 W$ represents the difference between the treated unit and the weighted combination of other units on pre-intervention characteristics
- The goal is to select the values of weights W that minimize this difference across all characteristics

Abadie, Diamond, and Hainmueller (2010) minimize W as follows:

$$\sum_{m=1}^{k} v_m (X_{1m} - X_{0m} W)^2,$$

where m denotes a pre-intervention variable, and v_m represents the relative importance assigned to that variable

The idea is that we want to minimize the distance between the treated unit and the units in the donor pool, but we care more about minimizing the distance for the variables that are most predictive of the outcome

How do we choose v_m?

- v_m denotes the importance of a variable m
- We care about minimizing the distance for some variables more than others
- e.g. If investment rate is strongly predictive of GDP, we want to minimize that distance between our synthetic control and West Germany a lot
- Typically we select v_m to minimize the squared prediction error:

$$\sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^*(v) Y_{jt} \right)^2$$

Once we have the set of weights W^*, we can estimate the effect of the treatment as follows:

$$Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

- The difference between the observed post-treatment outcomes Y_{1t} for the treatment unit and that of the weighted combination of outcomes from the donor units

How does one make inferences about the effect?

1. Run placebo intervention on the treatment unit at an earlier period in time
   - We shouldn't observe an effect for the treated unit before the intervention occurred
2. Run placebo interventions on all donor units at the time of the actual intervention (i.e. do a synthetic control for all donors)
   - Compare the effect obtained for the treatment unit to the placebo "effects" we observe for each of the donors
   - Sort of like randomization inference

Comparative Politics and the Synthetic Control Method

Alberto Abadie Harvard University and NBER
Alexis Diamond International Finance Corporation
Jens Hainmueller Stanford University

In recent years, a widespread consensus has emerged about the necessity of establishing bridges between quantitative and qualitative approaches to empirical research in political science. In this article, we discuss the use of the synthetic control method as a way to bridge the quantitative/qualitative divide in comparative politics. The synthetic control method provides a systematic way to choose comparison units in comparative case studies. This systematization opens the door to precise quantitative inference in small-sample comparative studies, without precluding the application of qualitative approaches. Borrowing the expression from Sidney Tarrow, the synthetic control method allows researchers to put "qualitative flesh on quantitative bones." We illustrate the main ideas behind the synthetic control method by estimating the economic impact of the 1990 German reunification on West Germany.

What was the effect of German re-unification on the economy in West Germany?

1. Use 16 OECD countries
2. Pre-unification variables: per capita GDP, inflation rate, industry share of value added, investment rate, schooling, trade openness

Weights obtained by minimizing difference between West German variables and those of donor countries

Table 1 Synthetic and Regression Weights for West Germany

Country	Synthetic Control Weight	Regression Weight	Country	Synthetic Control Weight	Regression Weight
Australia	0	0.12	Netherlands	0.09	0.14
Austria	0.42	0.26	New Zealand	0	0.12
Belgium	0	0	Norway	0	0.04
Denmark	0	0.08	Portugal	0	−0.08
France	0	0.04	Spain	0	−0.01
Greece	0	−0.09	Switzerland	0.11	0.05
Italy	0	−0.05	United Kingdom	0	0.06
Japan	0.16	0.19	United States	0.22	0.13

Notes: The synthetic weight is the country weight assigned by the synthetic control method. The regression weight is the weight assigned by linear regression. See text for details.

Weighted average of the synthetic control on pre-treatment characteristics

TABLE 2 Economic Growth Predictor Means before German Reunification

	West Germany	Synthetic West Germany	OECD Sample
GDP per capita	15808.9	15802.2	8021.1
Trade openness	56.8	56.9	31.9
Inflation rate	2.6	3.5	7.4
Industry share	34.5	34.4	34.2
Schooling	55.5	55.2	44.1
Investment rate	27.0	27.0	25.9

Notes: GDP per capita, inflation rate, trade openness, and industry share are averaged for the 1981–90 period. Investment rate and schooling are averaged for the 1980–85 period. The last column reports a population-weighted average for the 16 OECD countries in the donor pool.

Weighted average of the outcome variable (GDP per capita)

FIGURE 1 Trends in per Capita GDP: West Germany versus Rest of the OECD Sample

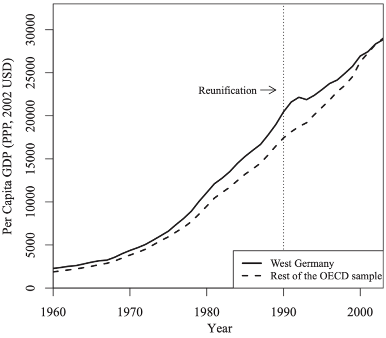
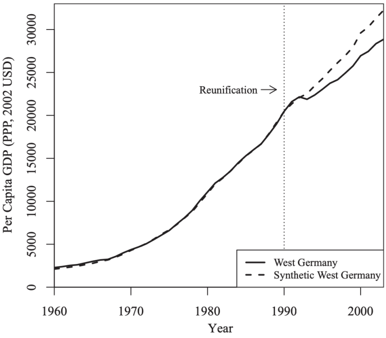


FIGURE 2 Trends in per Capita GDP: West Germany versus Synthetic West Germany



Difference between actual West Germany and synthetic

Figure 3 Per Capita GDP Gap between West Germany and Synthetic West Germany

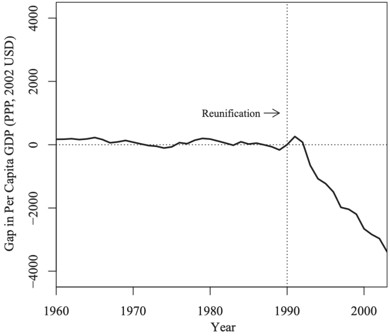
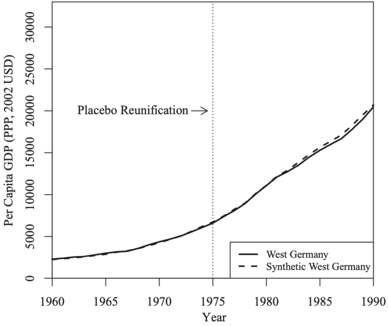


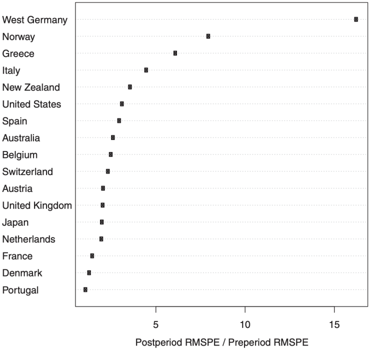
Figure 4 Placebo Reunification 1975–Trends in per Capita GDP: West Germany versus Synthetic West Germany



Placebo intervention ("reunification" in 1975) on the right

Placebo synthetic control on all other units individually

FIGURE 5 Ratio of Postreunification RMSPE to Prereunification RMSPE: West Germany and Control Countries



Much larger root mean square prediction error for West Germany compared to all other countries (highest out of 17)—can think of as a sort of p-value

Benefits of synthetic control

- Useful if want to know the effect of an event in a specific place
- Transparent about which units construct the synthetic case
- Very graphical, with figures that are simple to comprehend
- But takes a bit of work to understand how the weights are constructed, and how we make inferences about the estimated effect

Exercise

- Complete the exercise as usual on the course website