Section 10: Event Studies

Valentine Gilbert

April 21, 2022

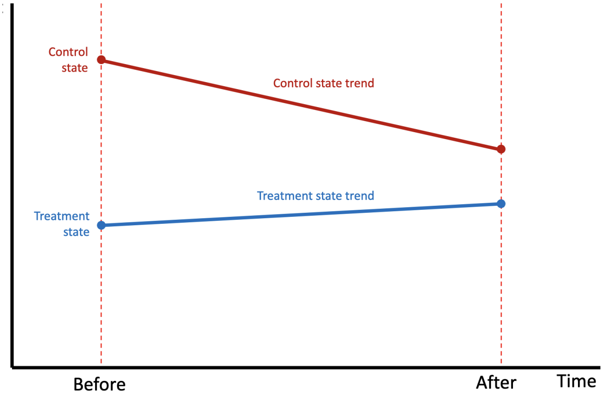
Overview

Simple DID: Regression

- The simplest diff-in-diff regression specification is:

$$Y_{it} = \beta_0 + \beta_1 D_i + \beta_2 P_t + \beta_3 (D_i \times P_t) + u_{it}$$

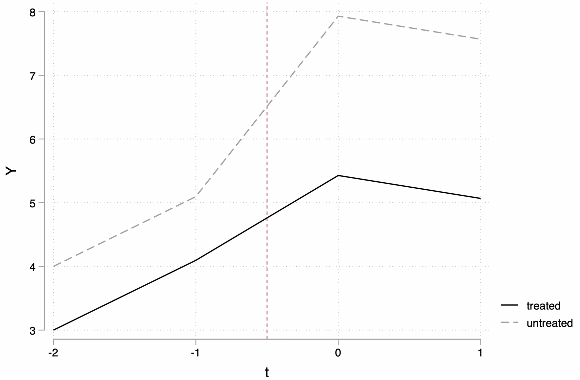
- Let's map the regression coefficients to the CEF:



Event Studies

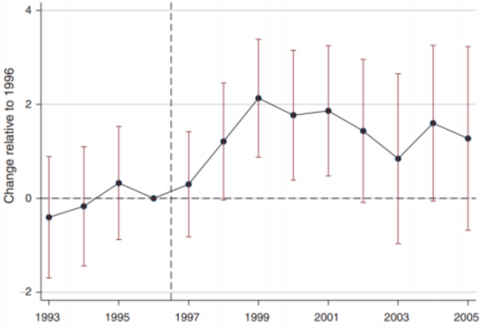
If we have two pre-periods and two post-periods we can estimate the following regression:

$$\begin{aligned} Y_{it} = &\alpha + \beta D_i + \gamma_1 1[t=-2] + \gamma_2 1[t=0] + \gamma_3 1[t=1] \\ &+ \delta_1 D_i \times 1[t=-2] + \delta_2 D_i \times 1[t=0] + \delta_3 D_i \times 1[t=1] \end{aligned}$$



Event Studies: Regression Specification

- Generalizing to the case with many periods and groups gives the event study specification:
$$Y_{it} = \sum_{\tau} \beta_t (D_i \times \mathbf{1}[\tau = t]) + \alpha_i + \gamma_t + \epsilon_{it}$$
- Plotting the coefficients and confidence intervals against the time periods gives the event study graph:



Event Studies: Regression Specification w/ Staggered Timing

- Finally, we can generalize to the case where treatment turns on at different times for different units:

$$Y_{it} = \sum_{\tau} \beta_{\tau}(D_i \times \mathbf{1}[\tau = t - t_i^*]) + \alpha_i + \gamma_t + \epsilon_{it}$$

Summary

Deshpande and Li (AEJ:EP 2019) – Who is screened out? Application costs and the targeting of disability programs

- Disability insurance programs in the US are large and expensive
  - SSDI: Cash benefits to 9 million workers in 2015
  - SSI: Cash benefits + Medicaid to 7 million low-income disabled Americans in 2015
  - Costly to verify disability status
- Potential trade-off between targeting and accessibility
  - Hassle costs could potentially improve targeting by screening out individuals who don't really need DI
  - Hassle costs could also potentially screen out those who need DI most
- This paper: Use variation in timing of SSA office closings to identify effect of application costs on targeting

Institutional background

Social Security Administration (SSA) administers:

- Social Security Disability Insurance (SSDI) – Requires a work history
- Supplemental Security Income (SSI) – Requires low income and assets

Both SSDI and SSI have the same medical requirements

As of 2019, $\sim$1,230 SSA field offices in the US

- One service is to help SSDI and SSI applicants in person and by phone
- "disability claims ... are particularly time intensive as employees help claimants complete detailed forms about medications, treatment, medical testing, work history, and daily activities"
- Office closings $\rightarrow$ increased application costs ($\uparrow$travel time, $\uparrow$congestion, etc.)

118 SSA office closings: Variation over time

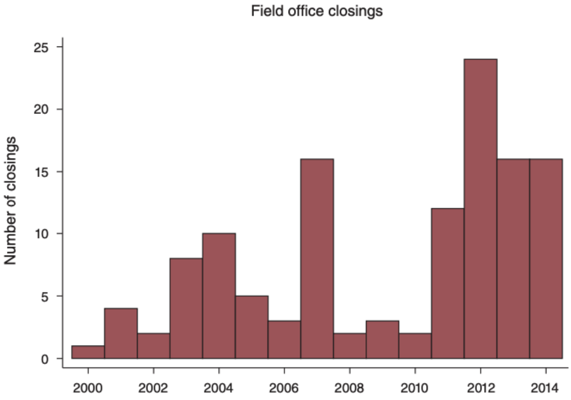


FIGURE 1. TIMING OF FIELD OFFICE CLOSINGS

“Closing”, “neighboring”, and “unaffected” ZIPs: Variation over space

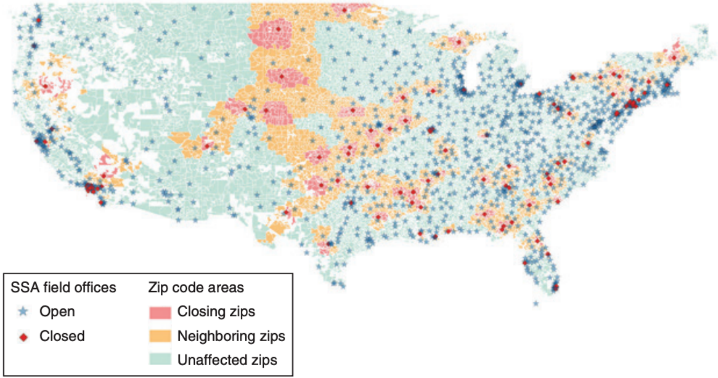


Figure 2. Map of Field Office Closings and Zip Code Classification in the United States

Empirical strategy

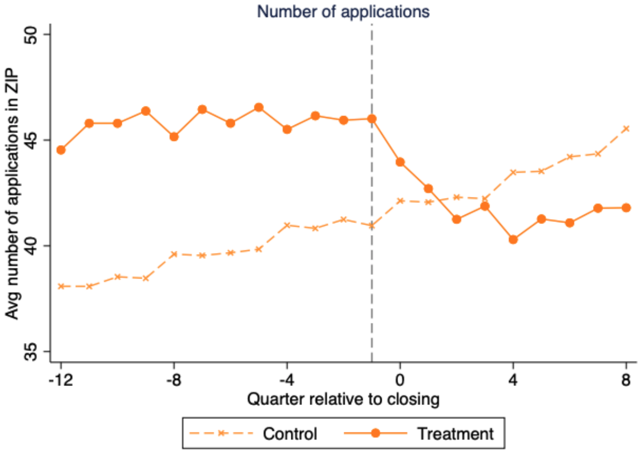
Authors find zip code characteristics predict if office *ever* closes but not *timing* of closing

$\Longrightarrow$ Use variation only from timing of closings to estimate effects

- Create separate datasets for each of the 118 closings
- In each dataset, zip codes experiencing the current closing are 'treated' and zip codes that experience closings > 2 years in future are 'control'
- Append all datasets and estimate event study

Raw data

Figure A.8: Raw Plots of Number of Applications in Control and Treatment ZIPs



Regression specifications

The authors estimate the following event study specification:

$$Y_{isct} = \alpha_i + \gamma_{st} + \delta_0 \, Treated_{ic} + \sum_{\tau} \gamma_\tau D_{ct}^\tau + \sum_{\tau \neq -2} \delta_\tau \left(Treated_{ic} \times D_{ct}^\tau \right) + \epsilon_{isct}$$

Diff-in-diff version:

$$\begin{aligned} Y_{isct} = \pi_i + \eta_{st} + \beta_0 \, Treated_{ic} + \sum_{\tau} \eta_\tau D_{ct}^\tau \\ + \beta \left(Treated_{ic} \times Post_{ct} \right) + \kappa \left(Treated_{ic} \times Zero_{ct} \right) + \epsilon_{isct} \end{aligned}$$

- Y_{isct}: Outcome for zip code i in state s for closing c in calendar-quarter t
- $Treated_{ic} = 1$ if zip code is affected by closing c
- $D_{ct}^\tau = \mathbf{1}[t - t_c^* = \tau]$ (i.e. indicates observation is τ quarters away from closing c)
- $Post_{ct} = \mathbf{1}[t > t_c^*]$, $Zero_{ct} = \mathbf{1}[t = t_c^*]$

Regression code: Stata

```
* Create lag variable (closing_c is the quarter of closing c)
gen lag = quarter - closing_c

* Keep lags -12 to 8 (following Deshpande and Li)
keep if lag >= -12 & lag <= 8

* Dummy out lag variable and create interactions
qui tab lag, gen(lag_)
foreach i of numlist 1/10 12/21 {  // NB: lag_11 = (lag == -2)
    gen treat_lag_`i' = treated*lag_`i'
}

* Create diff-in-diff variables
gen post = (quarter >  closing_c)
gen zero = (quarter == closing_c)
gen treat_post = treated*post
gen treat_zero = treated*zero

* Estimate event study
reg applicants i.zip i.state_quarter treated lag_* treat_lag_*, cluster(c)

* Estimate diff-in-diff
reg applicants i.zip i.state_quarter treated lag_* treat_post treat_zero, cluster(c)
```

Regression code: R

```r
# Create lag variable & subset data
stacked <- stacked %>%
  mutate(lag = quarter-closing_c) %>%
  filter(lag >= -12 & lag <= 8) %>%
  mutate(lag = factor(lag))

# Create lag dummies & create interactions
treat_lag <- stacked$lag
dummies <- model.matrix(~treat_lag+0)
dummies <- dummies[,-11]  # NB: column 11 corresponds to lag = -2
dummies <- dummies*stacked$treated
stacked <- cbind(stacked,dummies)

# Create DID variables
stacked <- stacked %>%
  mutate(
    post = quarter > closing_c,
    zero = quarter == closing_c,
    treat_post = treated*post,
    treat_zero = treated*zero
  )

# Estimate event study
lagvars <- collapse(colnames(dummies), collapse = "+")
es_formula <- paste("applicants ~ lag + treated +", lagvars, "| zip + state_quarter | 0 | c")
event_study <- felm(as.formula(es_formula), stacked) %>% tidy()

# Estimate DID
did <- felm(applicants ~ lag + treated + treat_post + treat_zero | zip + state_quarter | 0 | c, stacked) %>% tidy()

```

Event study: Disability applicants and recipients

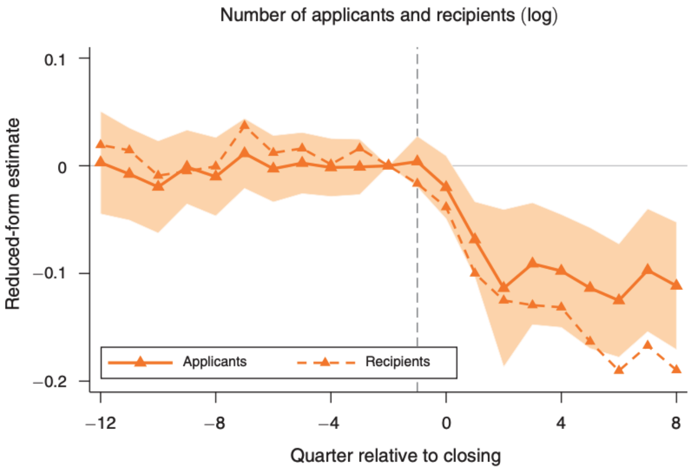


FIGURE 3. EFFECT OF CLOSINGS ON NUMBER OF DISABILITY APPLICATIONS AND ALLOWANCES

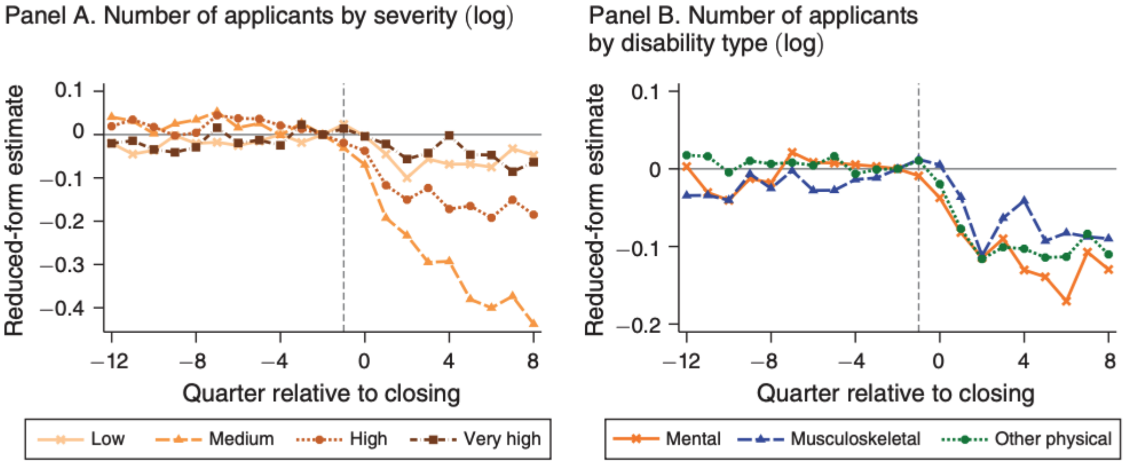
Implications for targeting

SSA office closings had a greater impact on log recipients than log applicants

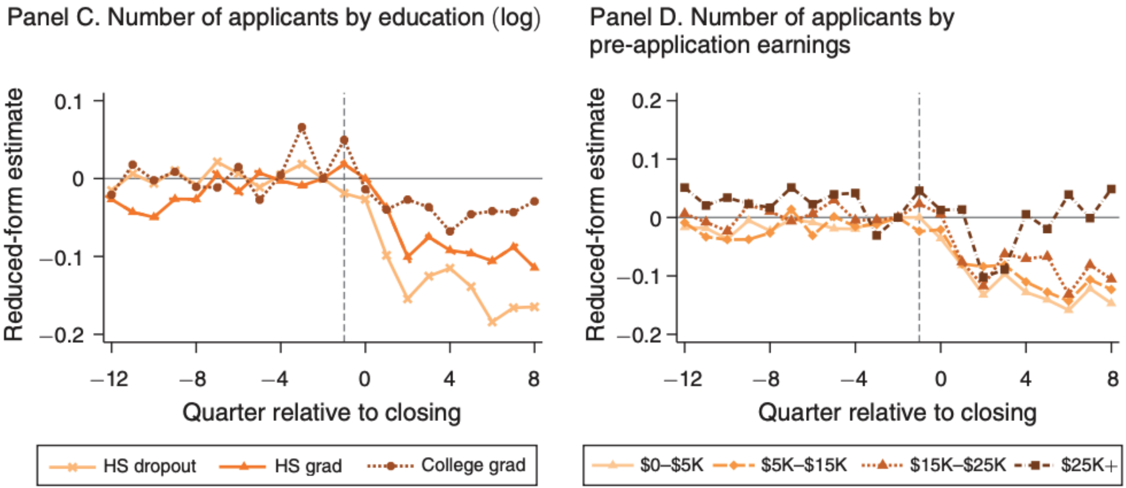
- -.1 log applicants vs -.16 log recipients in the post period
- **Question:** How should I interpret this?
- **Answer:** SSA office closings led to a greater % decrease in disability recipients than applicants
  $\implies$ increased application costs disproportionately deterred applicants who would have been approved

Deshpande and Li go on to estimate heterogeneity of effects along different dimensions

Event study heterogeneity: Health status



Event study heterogeneity: Socioeconomic status



Recap

So far, we've seen that SSA office closings:

- Decrease number of disability applicants and recipients
- Disproportionately screen out applicants who would be approved
- Disproportionately screen out applicants with medium/high severity disabilities
- Disproportionately screen out lower income and less educated applicants

Authors go on to ask how exactly do closings increase application costs

Mechanisms

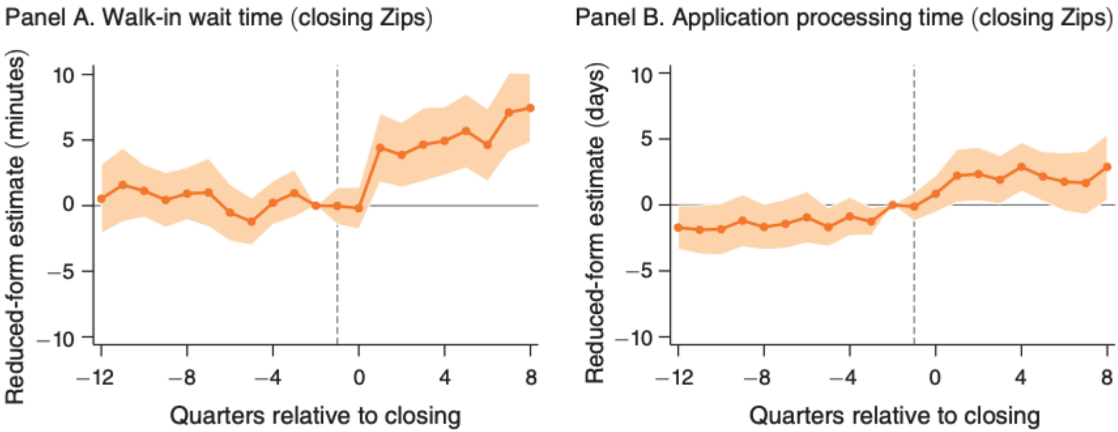
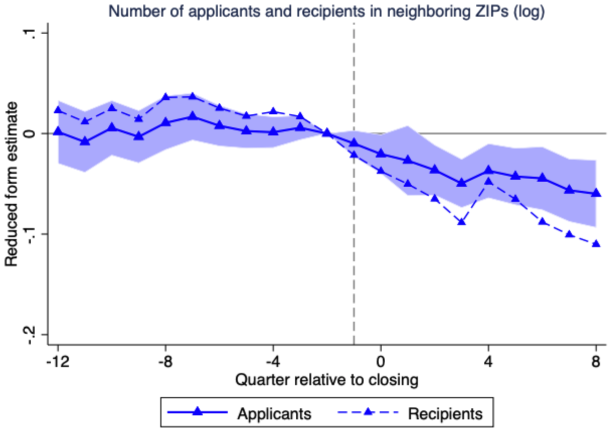


Figure 6. Effect of Closings on Measures of Field Office Congestion

Spillovers

Figure A.11: Effect of Closings on Number of Disability Applications and Allowances for Neighboring ZIPs

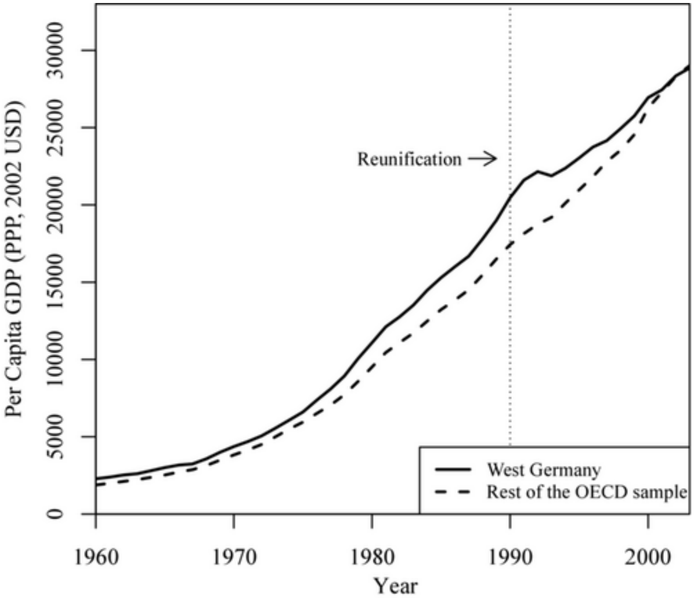


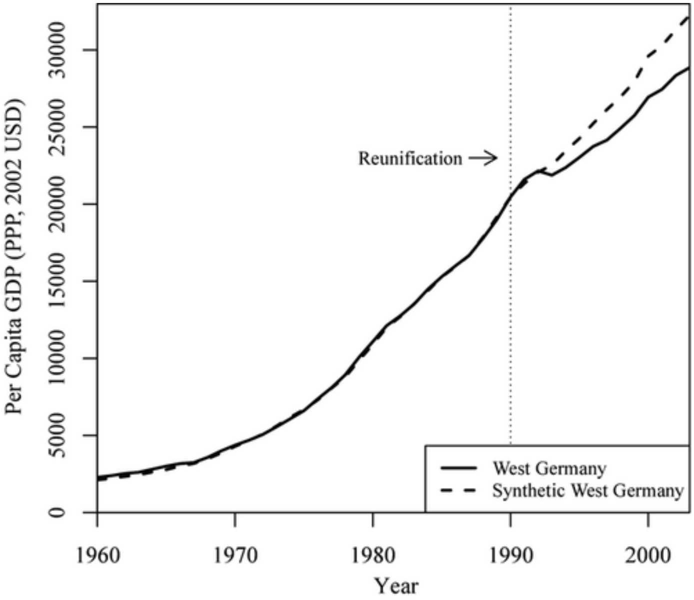
Takeaways

- Hassle costs decrease efficiency but can theoretically improve targeting
- In the case of disability applications, hassle costs decrease both efficiency and targeting:
  - Hassles disproportionately screen out applicants with moderately severe disabilities
  - Hassles disproportionately screen out applicants with less education and less income
- SSA office closings lead to congestion in neighboring Zip codes $\rightarrow$ hassles spill over

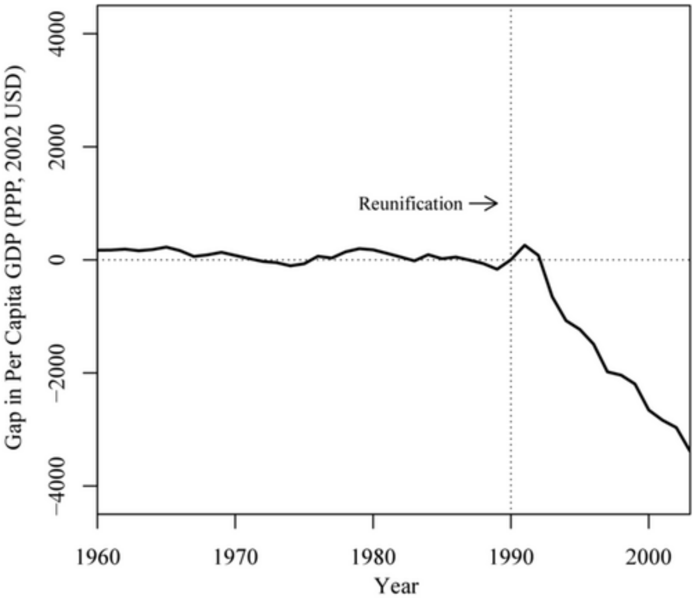
Synthetic Controls: Running Example

- After the fall of the Berlin Wall, East and West Germany were reunified in 1990. At the time of reunification, West Germany had three times the GDP per capita of East Germany.
- Abadie, Diamond, and Hainmueller (2014) use synthetic controls to estimate the effect of reunification on the West German economy
  - Find a weighted average of other economies that match West Germany on various characteristics in the pre-unification period
  - Use the weighted average of those economies' per capita GDP in the post-unification period as the counterfactual
- Hard to think of another method for quantitatively studying the effect of such a singular event


Per Capita GDP (PPP, 2002 USD)
0
5000
10000
15000
20000
25000
30000
1960
1970
1980
1990
2000
Year
Reunification
West Germany
Rest of the OECD sample



Per Capita GDP (PPP, 2002 USD)
0
5000
10000
15000
20000
25000
30000
1960
1970
1980
1990
2000
Year
Reunification
West Germany
Synthetic West Germany



Gap in Per Capita GDP (PPP, 2002 USD)
4000
2000
0
−2000
−4000
Reunification
1960
1970
1980
1990
2000
Year

Synthetic Controls: Weights

- So the synthetic control is a weighted average of comparison units
- But where do the weights come from?
- The weights are chosen to minimize the distance between the treated unit and the synthetic control unit on various covariates

$$\min_{W} \quad ||\boldsymbol{X_1} - \boldsymbol{X_0}\boldsymbol{W}||$$

- Abadie, Diamond, and Hainmueller (2014) match on GDP per capita, investment rate, trade openness, schooling, inflation rate, and industry share

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

Synthetic Controls: Weights

- But if we're matching on many covariates, we may not be able to improve the match on one without making the match on another covariate worse. How do we balance this trade-off?
- This is where the diagonal matrix $\boldsymbol{V}$ comes in:

$$
\begin{aligned}
||\boldsymbol{X}_1 - \boldsymbol{X}_0 \boldsymbol{W}|| &= \sqrt{(\boldsymbol{X}_1 - \boldsymbol{X}_0 \boldsymbol{W})' \boldsymbol{V} (\boldsymbol{X}_1 - \boldsymbol{X}_0 \boldsymbol{W})} \\
&= \sqrt{\sum_{m=1}^{k} v_m (X_{1m} - X_{0m} W)^2}
\end{aligned}
$$

- $\boldsymbol{V}$ is another set of weights that tells us how important it is to match on each covariate (i.e. how predictive each covariate is of the outcome)
- For example, it's important to get a really good match on per capita GDP in the pre-unification period, even if that means sacrificing the match quality on per capita consumption of bratwurst

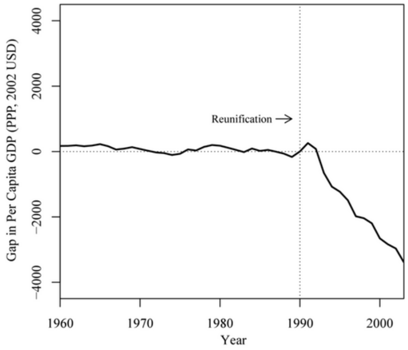
Synthetic Controls: Choosing V

- To recap, our counterfactual is a weighted average of comparison places and those weights depend on a *second* set of weights, $\boldsymbol{V}$
- Here are three approaches to choosing $\boldsymbol{V}$:
  - Make each weight inversely proportional to the variance of the covariate, so they all have unit variance
  - Choose $\boldsymbol{V}$ such that $W(\boldsymbol{V})$ minimizes the pre-period prediction error (i.e. choose $\boldsymbol{V}$ that minimizes in-sample prediction error)
  - Choose $\boldsymbol{V}$ such that $W(\boldsymbol{V})$ calculated in the first half of the pre-period minimizes the prediction error in the latter half of the pre-period (i.e. choose $\boldsymbol{V}$ that minimizes out-of-sample prediction error)
- Abadie, Diamond, and Hainmueller take the third approach, using the years 1971-1980 as a training period and 1981-1990 as the validation period and choosing $\boldsymbol{V}$ that minimizes

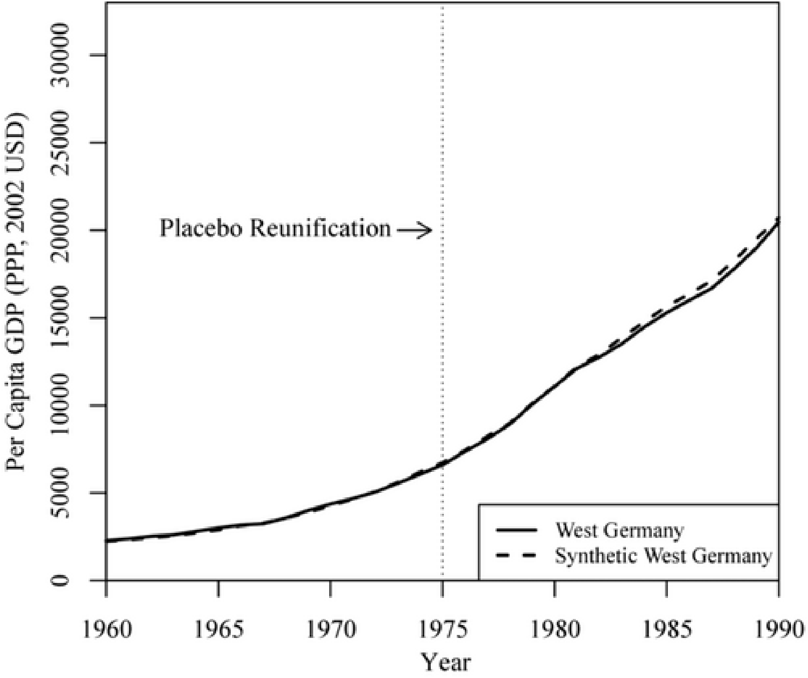
$$\text{RMSPE} = \left(\frac{1}{10}\sum_{t=1981}^{1990}\left(GDP_{1t} - \sum_{j=2}^{J+1} w_j^*(\boldsymbol{V})GDP_{jt}\right)^2\right)^{1/2}$$

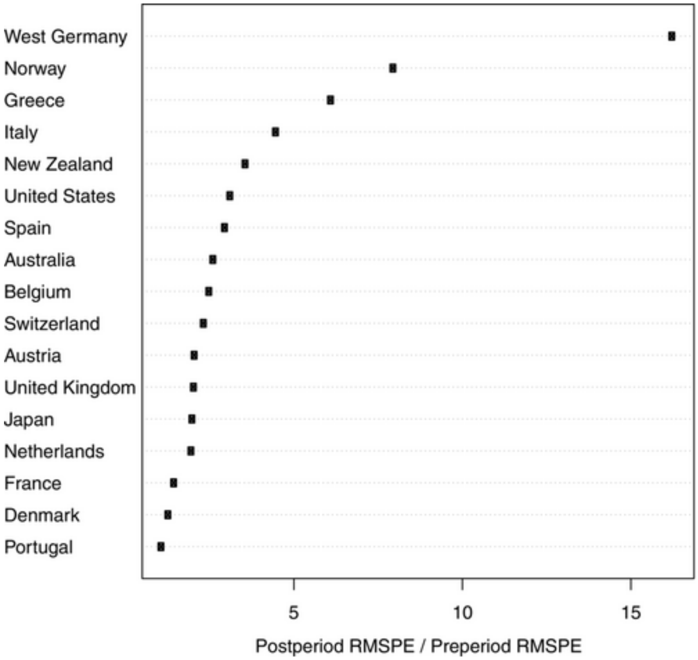
Synthetic Controls: Inference

- So we chose $\boldsymbol{V}$, which gave us our weights, which gave us a synthetic control group, which gave us this nice event study:

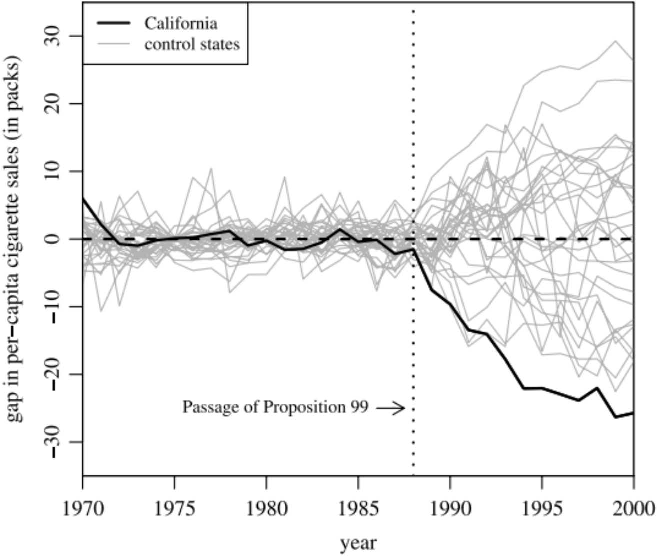


- How do we know if the effect we see in the post-unification period is significant?
- Randomization inference! (AKA placebo studies)


Per Capita GDP (PPP, 2002 USD)
0
5000
10000
15000
20000
25000
30000
Placebo Reunification
West Germany
Synthetic West Germany
1960
1965
1970
1975
1980
1985
1990
Year


West Germany
Norway
Greece
Italy
New Zealand
United States
Spain
Australia
Belgium
Switzerland
Austria
United Kingdom
Japan
Netherlands
France
Denmark
Portugal
5
10
15
Postperiod RMSPE / Preperiod RMSPE

They don't present this kind of plot for their study of German reunification (why?), but plotting all of the event studies can be a good way of presenting results:



Synthetic Controls vs. Regression

- An alternative approach to estimating the counterfactual outcome is to use regression
- Specifically, we could take the donor pool and regress the outcome of interest in each post-intervention period on the observed covariates. We could then use the regression coefficients to predict a counterfactual for the treated unit.
- The regression-based counterfactual is also a weighted average of outcomes observed for comparison units, where the weights sum to one
- But the weights don't necessarily lie within [0,1], and the method is much less transparent

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

Synthetic Controls: Further Reading

- This article is forthcoming in the JEL and looks like a useful guide to SCM: Abadie (2019 JEL)
- Gives overview of SCM, discusses contextual and data requirements, and discusses extensions