

6. Synthetic control methods

LPO 8852: Regression II

Sean P. Corcoran

Synthetic control method - introduction

Synthetic control methods are often used when there is only *one* treated unit observed before and after treatment and no clear comparison unit.

- A context where one might like to use difference-in-differences but potential comparison units are quite different from the treated unit, such that the parallel trends assumption fails to hold.
- The treated unit is often at a high level of aggregation (e.g., country, region, state) but could be a smaller unit (e.g., school or firm).
- Abadie and Gardeazabal (2003) is the classic reference—on the impact of terrorism on economic activity.
- Abadie et al. (2010) elaborate on the methods in the context of an anti-smoking law in California.
- Abadie (2021) is an excellent survey—highly recommended.

Motivating example

A classic paper by Card (1990) looks at the effect of the Mariel Boatlift from Cuba in 1980 on the Miami labor market. Cuban immigrants increased the size of the local labor force by 7%. What effect did this have on the employment of less-skilled native workers?

- Used Current Population Survey data on the employment status of native born workers. One treated labor market (Miami).
- Estimated a standard difference-in-differences using similar workers in Atlanta, Los Angeles, Houston, and Tampa.
- The selection of comparisons was arguably *ad hoc*.
- This study famously found no impact of immigration in Miami on local unemployment.

Synthetic control

Synthetic control methods optimally choose a set of weights that—when applied to a group of corresponding untreated units—provides a counterfactual path for the unit that received the treatment.

- This weighted group is the “synthetic unit” and stands in for what would have happened to the aggregate treated unit had the treatment not occurred.

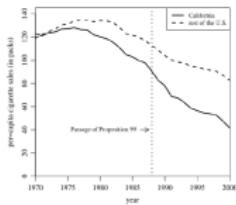


Figure 1. Trends in per capita cigarette sales: California vs. the rest of the United States.



Figure 2. Trends in per capita cigarette sales: California vs. synthetic California.

Figures 1 and 2 from Abadie et. al. (2010)

Synthetic control

Advantages:

- Doesn't require a large sample of treated and untreated cases.
- Selection of comparison units—and their exact weighting—is explicit.
- Doesn't extrapolate as is done in regression; estimates are always a weighted average of other non-treated units.
- Only need pre-treatment data to construct weights.

Disadvantage:

- Requires a sufficient number of pre-treatment observations to identify a "good" synthetic control.

Formalization of synthetic control method

Formalization

- Y_{jt} is the outcome of interest for unit j in period t
- There are $J + 1$ units and unit 1 is treated. The other J units are the **donor pool**.
- There are T periods, $1, \dots, T_0$ before treatment and $T_0 + 1, \dots, T$ after
- We may also observe predictors of Y_{jt} that are time-varying: X_{jkt} ($k = 1, \dots, K$)

Potential outcomes

- Can think about potential outcomes $Y_{jt}(1)$ and $Y_{jt}(0)$ so that the treatment effect for unit j is $\tau_{jt} = Y_{jt}(1) - Y_{jt}(0)$. Note τ is *time varying*.
- As always, the challenge is to estimate $Y_{jt}(0)$ for the treated case in the treated period(s). That is, how would Y_{jt} have evolved in the absence of treatment?
- **Synthetic control:** finding a weighted combination of units in the donor pool to approximate $Y_{jt}(0)$:

$$\hat{Y}_{jt}(0) = \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

where w_j^* is a set of optimally chosen weights.

Potential outcomes

- So then the estimated treatment effect for unit 1 in time t is:

$$\hat{\tau}_{1t} = Y_{1t}(1) - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

- The goal is to identify a weighted combination of units in the donor pool that approximates $Y_{1t}(0)$ well in the *pre-treatment* period. The assumption is that if this synthetic group tracked the treated observation in the pre-period, it would continue to do so in the post-period, in the absence of treatment.
- How do we find this donor pool? Can use pre-treatment values of Y and/or strong predictors of Y .

Finding weights

So how do we obtain these weights? What criteria do we use?

- Usually, weights are constrained to be non-negative and sum to one.
- One possibility would be to just use equal weights $w_j = 1/J$:

$$\hat{\tau}_{1t} = Y_{1t}(1) - \frac{1}{J} \sum_{j=2}^{J+1} Y_{jt}$$

- Or a population-weighted version where w_j^{pop} is the size of unit j as a fraction of the total donor pool:

$$\hat{\tau}_{1t} = Y_{1t}(1) - \sum_{j=2}^{J+1} w_j^{pop} Y_{jt}$$

Not surprisingly, these often do not perform that well, and there is no reason to think these would provide a good counterfactual.

Finding weights

Abadie and Gardezabal (2003) and Abadie et al. (2010) propose choosing weights so that the synthetic control best approximates the treated unit before the intervention. Their method involves two types of weights:

- **Unit weights** (w_j): how much each donor unit contributes to the synthetic control.
- **Predictor or variable weights** (v_k): how important each predictor variable is for matching on pre-treatment outcomes (Y) (i.e., minimize the MSPE for pre-treatment outcomes).

Goal: choose w_j and v_k to minimize the distance between the treated unit and the synthetic unit in the pre-period on all predictor variables:

$$||\mathbf{X}_1 - \mathbf{X}_0 \mathbf{W}|| = \left(\sum_{h=1}^k v_h (X_{h1} - w_2 X_{h2} - \dots - w_{J+1} X_{hJ+1})^2 \right)^{1/2}$$

Finding weights

Intuitively, this is a nested (two-step) procedure:

- For a given set of v , find the optimal w (subject to $0 \leq w_j \leq 1 \forall j$)
- Choose v to maximize pre-treatment fit on Y

The process is iterative, changing v to obtain better pre-treatment fit. (The optimal w depends on v).

According to Abadie (2021), *sparsity* is typical of the synthetic control method—i.e., the resulting weights tend to be concentrated on a small number of units. This is a feature, since it makes the synthetic group more transparent to the reader.

Synthetic control: assumptions

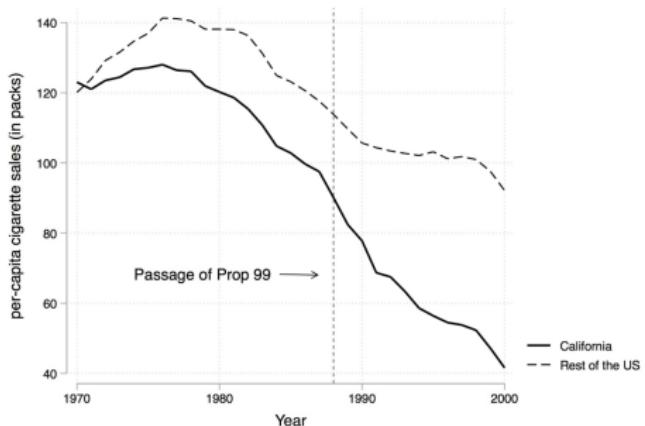
What assumptions does the SCM rely upon for causal inference?

- ① There exists a “synthetic control” (with nonnegative weights summing to 1) that would have followed the treated unit’s path in the absence of treatment. Analogous to parallel trends.
- ② “Convex hull” assumption: the unit’s pre-treatment characteristics lie within the convex hull of the donor units. In other words, treated unit cannot be *too different* such that one cannot find a valid counterfactual from the donor units.
- ③ Stability of the relationship between predictors and outcomes: the relationship would be the same in the absence of treatment.
- ④ No interference (SUTVA)
- ⑤ No anticipatory effects
- ⑥ Good pre-treatment fit

Classic example: Abadie et al. (2010)

Example: Abadie et al. (2010)

The impact of Proposition 99 on per-capita cigarette sales. Prop 99 increased cigarette taxes and funded other anti-smoking initiatives.



Example: Abadie et al. (2010)

Per-capita cigarette sales for California and “synthetic California”

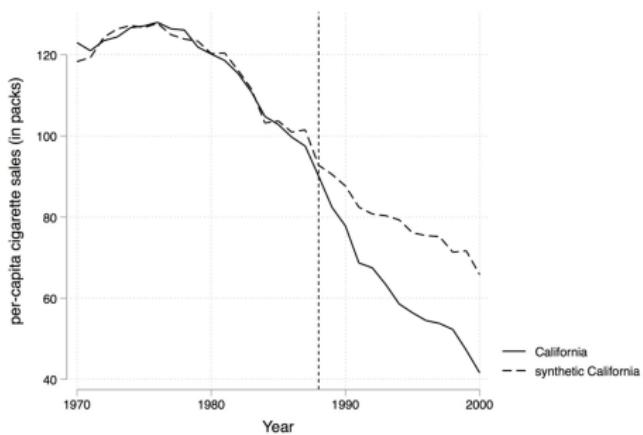


Figure 10.2: California cigarette sales vs synthetic California

Example: Abadie et al. (2010)

Variables used to find optimal weights—includes both predictors of cigarette sales *and* pre-treatment values of cigarette sales:

Table 10.1: Balance table

Variables	Real California	Synthetic Calif.	Avg. of 38 Control States
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15–24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	120.43	136.58
Cigarette sales per capita 1975	127.10	126.99	132.81

All variables except lagged cigarette sales are averaged for the 1980–1988 period. Beer consumption is averaged 1984–1988.

Example: Abadie et al. (2010)

Another nice way to show pre- and post-treatment gap in the outcome:

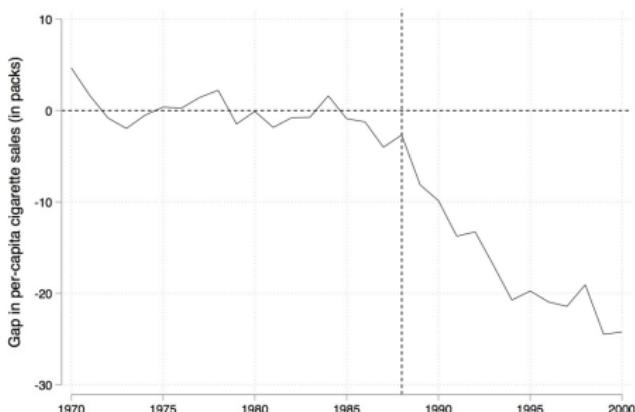


Figure 10.3: Gap in cigarette sales for estimation pre and post treatment

Statistical inference with the synthetic control method

Inference

The graph is compelling, but how can we say whether the post-treatment difference is *significant* or not?

- Inference is atypical here since we only have one observation (per year) on the treated and synthetic control group.
- Similarly, cannot formally test for baseline differences in Table 10.1 above since there is only one observation in each group.
- Abadie et al. (2010) propose **placebo-based** inference:
 - ▶ It's possible that the SCM can yield differences even when there is no treatment effect.
 - ▶ Apply the SCM to *every* unit in the donor pool, treating that unit as the "treated" case, and calculate the (placebo) treatment effect.
 - ▶ How *unusual* is the (actual) treatment effect relative to these?

Example: Abadie et al. (2010)

All placebo cases overlaid with the actual treatment case (in bold):

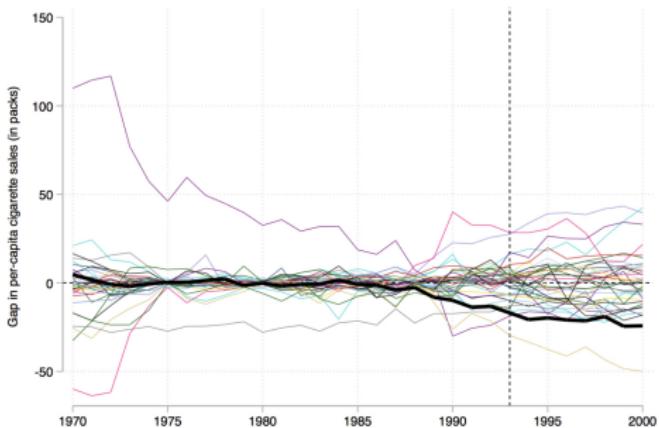
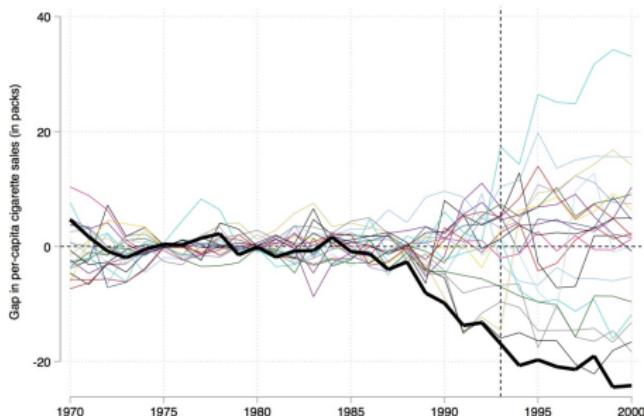


Figure 10.4: Placebo distribution using all units as donor pool

Example: Abadie et al. (2010)

Dropping units with very poor fit pre-treatment can help this graph:



Inference

In addition to visual evidence, we can calculate a test statistic:

- ① Apply SCM to each unit in the donor pool
- ② Calculate the mean square prediction error (MSPE) in the *pre-treatment* period (i.e., how well does the donor pool track the “treated” case in the pre-period?)
- ③ Calculate the MSPE in the *post-treatment* period (i.e., how closely does the donor pool track the “treated” case in the post-period? If there is an effect, these should diverge).
- ④ Compute the *ratio* of post- to pre-MSPE.
- ⑤ Sort this ratio from highest to lowest.
- ⑥ Calculate the rank of the (actual) treatment unit ratio in this distribution and calculate a “*p*-value” (Rank/Total).

Example applications

Applications: education and health

- Shores, Candelaria & Kabourek (2023): effects of school finance reforms. Reform states contrasted with a weighted average of non-reform states.
- Gutierrez, Weinberger, & Engberg (2016): Gates Foundation Intensive Partnerships for Effective Teaching program.
- Bifulco, Rubenstein, & Sohn (2017): "Say Yes to Education" promise scholarship in Syracuse.
- Dave et al. (2021): impact of Trump rally in Tulsa on COVID-19 spread.
- Trejo et al. (2021): academic and psychosocial effects of Flint water crises (e.g., anxiety and depression)

Bifulco et al. (2017)

"Say Yes to Education" was a promise scholarship program implemented in Syracuse NY in 2008. It offers full tuition scholarships for public and private universities, coupled with wraparound support services in schools.

- They consider two donor pools, a restricted one (judged similar to Syracuse in important ways) and a comprehensive one (all urban or suburban districts in NYS). See Table 2.
- They try six alternative combinations of pre-treatment year outcome means to find weights. See Table 1. They also use average shares FRPL, African-American, and Hispanic (over the whole pre-period).
- There are two different outcomes: enrollment and graduation rates. Weights can and will differ depending on the outcome.

Alternative combinations of pre-treatment years to find weights (Table 1):

Table 1. Alternative Specifications of Pretreatment Years.

Specification	Description
1	First and last year of pretreatment periods
2	First, middle, and last year of pretreatment periods
3	Middle and last year of pretreatment periods
4	Last pretreatment year and the average of outcomes in all other pretreatment years
5	Each pretreatment year
6	Each year from the middle to the end of the pretreatment periods

Table 2. Assignment of Weights (Enrollment Analysis).

District Name	Assigned Weights					
	Specif. 1	Specif. 2	Specif. 3	Specif. 4	Specif. 5	Specif. 6
Panel A: Restricted donor pool						
Albany CSD	.000	.021	.000	.000	.005	.000
Brentwood UFSD	.000	.000	.000	.000	.116	.129
Buffalo CSD	.000	.078	.022	.034	.029	.000
Niagara Falls CSD	.484	.288	.498	.499	.411	.404
Rochester CSD	.502	.406	.479	.467	.438	.467
Utica CSD	.014	.207	.001	.000	.000	.000
Panel B: Comprehensive donor pool						
Albany CSD	.000	.000	.000	.000	.005	.000
Brentwood UFSD	.000	.000	.000	.000	.116	.134
Buffalo CSD	.117	.174	.091	.065	.029	.004
Elmira CSD	.000	.000	.307	.148	.000	.000
Hopevale UFSD	.101	.197	.053	.021	.000	.000
Mount Vernon CSD	.000	.061	.000	.000	.000	.000
Niagara Falls CSD	.248	.000	.124	.324	.411	.401
Rochester CSD	.386	.341	.425	.442	.438	.461
Smythtown CSD	.149	.156	.000	.000	.000	.000
Utica CSD	.000	.069	.000	.000	.000	.000

Note. Specif., CSD, and UFSD denote "specification," "city school district," and "union free school district," respectively. Districts that do not appear in the table do not receive positive weights equal to or greater than 0.001 in any of the specifications. Restricted donor pool includes 22 districts, whereas comprehensive donor pool includes 275 districts.

Results: enrollment outcome (Figure 2)

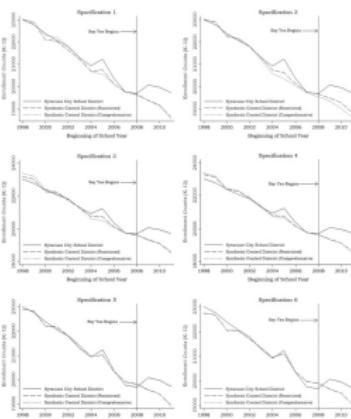


Figure 2. Trends in enrollment by model specifications. See Table 1 for description of pretreatment years included in each specification.

Bifulco et al. (2017)

On inference:

- Following Abadie et al. (2010), they use a permutation test to assess the likelihood that the effect found would occur by chance.
- They apply SCM for each unit in the donor pool and compare the “real” effect to these “placebo” effects.
- For each iteration of SCM, they compute RMSPE: the square root of the average of the squared prediction errors across years in the pre- and post-treatment periods. The ratio of these is a test statistic.
- They calculate the percentage of test statistics that are as large as that obtained for Syracuse. Interpreted as the probability of obtaining effect estimates this large if the true treatment effect were zero.

Results: enrollment effects

Table 3. Estimated Effects on K-12 Enrollments, RMSPE, and *p*-value.

Specification	Year 1	Year 2	Year 3	Year 4	RMSPE	<i>p</i> -value
Panel A: Restricted donor pool						
Specification 1	30	704	789	1,110	214.16	.087
Specification 2	24	576	676	889	220.09	.217
Specification 3	69	739	840	1,166	209.04	.044
Specification 4	64	730	839	1,165	227.19	.130
Specification 5	-114	500	560	795	114.27	.304
Specification 6	-216	405	445	672	161.23	.130
Panel B: Comprehensive donor pool						
Specification 1	87	693	859	1,132	252.35	.076
Specification 2	147	715	920	1,117	318.44	.243
Specification 3	76	702	845	1,075	323.72	.098
Specification 4	67	713	840	1,121	280.36	.091
Specification 5	-114	500	560	795	114.27	.562
Specification 6	-213	404	445	668	156.37	.120

Note. Restricted donor pool includes 22 districts, whereas comprehensive donor pool includes 275 districts. Years 1–4 correspond to the effect estimates. *p*-value implies a probability of getting a post/pretreatment RMSPE ratio as large as the post/pretreatment RMSPE ratio of Syracuse if one assigns the treatment at random in the data. Specifications are the same as in Table 1. Pretreatment period includes years 1998–2007. All models are run with percent Black, percent Hispanic, and percent free lunch eligible as covariates. RMSPE = root mean squared prediction error.

Synthetic control in Stata

Synthetic control in Stata: synth2

Use the synth2 package to get much of what you want. Requires the original Abadie et al. synth package (and possibly mat2txt). Note synth2 only works in recent versions of Stata (16+). If this package doesn't work for you, you can use synth and synth_runner for inference tools. See the in-class exercise for an example. Syntax:

`synth2 depvar predictorvars, options`

Note: see Yan & Chen (2023) for more on synth2. Galiani & Quistorff (2017) describe synth_runner.

Synthetic control in Stata: synth2

Options:

- `trunit()`: ID of treatment unit
- `trperiod()`: time period when treatment occurs
- `mspeperiod()`: pre-treatment period used to identify synthetic control (MSPE = mean squared prediction error)
- `preperiod()`: pre-treatment time period (need not be the same as mspeperiod)
- `postperiod()`: post-treatment time period
- `xperiod()`: defines periods over which predictors are averaged, where applicable (need not be the same as mspeperiod).
- Others (see help menu)

synth2 output

The synth2 command does a lot of work for you. It provides:

- RMSPE in pre-period
- v-weights (predictors)
- Unit weights (units that make up the synthetic control)
- Predictor balance in the pre-period: treated, synthetic control, average control (all controls, unweighted)
- Post-treatment actual outcome, synthetic control outcome, "treatment effect" by period
- Graphs:
 - ▶ Predictor balance (before weighting)
 - ▶ Actual vs synthetic outcomes, pre and post periods
 - ▶ Treatment effect (diff between actual and synthetic)
 - ▶ v-weights, unit weights

In-class example: from the *Mixtape*

What was the impact of prison construction in Texas on the incarcerated population? In 1993, Gov. Richards funded a significant expansion of prisons in Texas in response to an earlier court ruling about prison over-crowding. Did these new prisons increase the incarcerated population, perhaps due to a curtailment of parole? Did expanded prison capacity have a disproportionate impact on Black males?

- See annotated do file on Github using synth2
- Use other states as a synthetic control, estimate treatment effect
- Placebo inference and leave-one-out robustness check

The above .do file also shows the use of the older synth and synth_runner. Don't use these unless you cannot run synth2.

Placebo inference using synth2

synth2 also has the following options:

- “In-space” placebo test (the standard placebo test used for inference, described earlier; from Abadie et al., 2010)
- “In-time” placebo test (sets the treatment year to an earlier “fake” treatment year; from Abadie et al., 2015)
- A “leave-one-out” (LOO) robustness test: iteratively removes one selected donor at a time (from Abadie et al., 2015)

See example output below from the in-class exercise, and also Yan & Chen (2023).

Placebo inference using synth2

Interpreting “ p -values” from in-space placebo test:

- **two-sided p -value:** the proportion of times that the *absolute values* of the placebo effects are greater than or equal to the absolute value of the estimated treatment effect.
- **right-sided p -value:** the proportion of times that the placebo effects are greater than or equal to the estimated treatment effect.
- **left-sided p -value:** the proportion of times that the placebo effects are smaller than or equal to the estimated treatment effect.

Note that these can be calculated for every post-treatment period. For an “overall” measure of significance, synth2 also reports the likelihood of obtaining a post/pretreatment MSPE as large as the focal unit’s.

In-class example: synth2 output

```
synth2 bmrates bmrates(1988) bmrates(1991) bmrates(1992) ///
>      alcohol(1990) aidscapita(1990) aidscapita(1991) income ur ///
>      poverty black(1990) black(1991) black(1992) perc1519(1990), ///
>      trunit(48) tperiod(1993) mspeperiod(1985(1)1992) ///
>      preperiod(1985(1)1992) postperiod(1993(1)2000) ///
>      xpriod(1985(1)1992) /*nested allop*/ ///
>      savegraph(set1, replace)
```

Fitting results in the pretreatment periods:

Treated Unit	:	Texas	Treatment Time	:	1993
Number of Control Units	=	58	Root Mean Squared Error	=	67.42838
Number of Covariates	=	13	R-squared	=	0.96543

Covariate balance in the pretreatment periods:

Covariate	V.weight	Treated	Synthetic Control Value	Average Control Bias	Average Control Value	Bias
bmrates(1988)	0.0028	1764.5443	1775.8312	0.64%	1981.8483	7.88%
bmrates(1991)	0.0013	2301.0017	2349.9358	-2.13%	2351.1634	2.18%
bmrates(1992)	0.0012	2673.3381	2585.7208	-3.28%	2449.8657	-8.36%
alcohol(1990)	0.0000	2.4700	2.4078	-2.52%	2.5840	1.38%
aidscapita(1990)	0.0009	13.5899	13.5291	-0.45%	8.7631	-35.52%
aidscapita(1991)	0.0007	14.8350	14.7733	-0.43%	10.1358	-31.74%
income	0.0000	16108.6158	16204.1307	0.59%	17080.3450	6.03%
ur	0.0000	7.4266	7.0739	-4.74%	6.1534	-17.14%
poverty	0.0000	17.2000	16.2830	-5.33%	13.3922	-22.14%
black(1990)	0.2166	16.1546	16.1601	0.03%	11.2884	-30.12%
black(1991)	0.6479	16.2998	16.2984	-0.01%	11.4188	-29.95%
black(1992)	0.1285	16.4595	16.4159	-0.27%	11.5657	-29.73%
perc1519(1990)	0.0000	7.7030	7.2955	-5.28%	7.1928	-6.62%

Note: "V.weight" is the optimal covariate weight in the diagonal of V matrix.

"Synthetic Control" is the weighted average of donor units with optimal weights.

"Average Control" is the simple average of all control units with equal weights.

In-class example: synth2 output

Optimal Unit Weights:

Unit	U.weight
Tennessee	0.1910
Oklahoma	0.1390
Wisconsin	0.1290
Montana	0.1100
WestVirginia	0.1090
DistrictofColumbia	0.1020
Arkansas	0.1010
Illinois	0.0670
NewYork	0.0440
Louisiana	0.0070

Note: The unit Alabama Alaska Arizona California Colorado Connecticut Delaware Florida Georgia Hawaii Idaho Indiana Iowa Kansas Kentucky Maine Maryland Massachusetts Michigan Minnesota Mississippi Missouri Nebraska Nevada NewHampshire NewJersey NewMexico NorthCarolina NorthDakota Ohio Oregon Pennsylvania RhodeIsland SouthCarolina SouthDakota Utah Vermont Virginia Washington Wyoming in the donor pool get a weight of 0.

In-class example: synth2 output

Prediction results in the posttreatment periods:

Time	Actual Outcome	Synthetic Outcome	Treatment Effect
1993	2770.3533	2710.5056	59.8477
1994	3748.4478	2899.9797	848.4680
1995	5089.6675	3043.0000	1966.6675
1996	4920.5317	3234.2971	1686.2346
1997	5049.5547	3337.7361	1711.8186
1998	5068.4648	3605.2271	1463.2378
1999	5083.7988	3752.2852	1331.5137
2000	5330.9468	3735.8499	1595.0969
Mean	4622.7207	3289.8601	1332.8606

Note: The average treatment effect over the posttreatment period is 1332.8606.

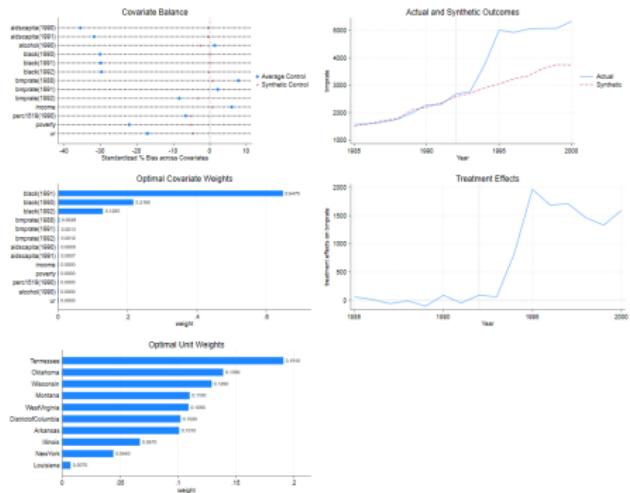
In-class example: synth2 output

Implementing placebo test using fake treatment unit Alabama..Alaska..Arizona..Arkansas..California..Colorado..Connecticut..Delaware..DistrictofColumbia..Florida..Georgia..Hawaii..Idaho..Illinois..Indiana..Iowa..Kansas..Kentucky..Louisiana..Maine..Maryland..Massachusetts..Michigan..Minnesota..Mississippi..Missouri..Montana..Nebraska..Nevada..NewHampshire..NewJersey..NewMexico..NewYork..NorthCarolina..NorthDakota..Ohio..Oklahoma..Oregon..Pennsylvania..RhodeIsland..SouthCarolina..SouthDakota..Tennessee..Utah..Vermont..Virginia..Washington..WestVirginia..Wisconsin..Wyoming...

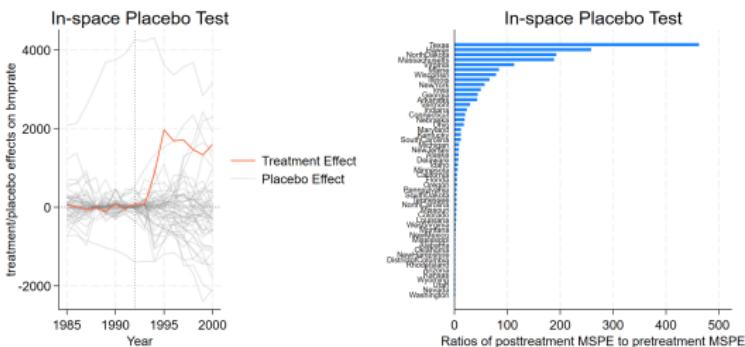
In-space placebo test results using fake treatment units:

Unit	Pre MSPE	Post MSPE	Post/Pre MSPE	Pre MSPE of Fake Unit/ Pre MSPE of Treated Unit
Texas	4546.5863	2.10e+06	462.5255	1.0000
Alabama	7145.3228	9508.9957	1.3308	1.5716
Alaska	3.22e+04	2.47e+05	7.6668	7.0818
Arizona	2.26e+05	2.08e+05	0.9166	49.8117
Arkansas	2704.1346	1.16e+05	42.8731	0.5948
California	2.44e+04	1.18e+05	4.8566	5.3630
Colorado	5.23e+04	1.65e+05	3.1663	11.4951
Connecticut	5.74e+04	1.16e+06	20.1676	12.6314
Delaware	6.34e+04	4.65e+05	7.3315	13.9528
DistrictofColumbia	1.08e+07	1.31e+07	1.2202	2378.0152
Florida	6048.6280	2.85e+04	4.7173	1.3304
Georgia	843.9226	3.70e+04	43.8205	0.1856
Hawai	8528.3743	2.26e+06	258.2551	1.8758
Idaho	1.42e+05	9.16e+05	6.4422	31.2631
Illinois	1556.5487	1.03e+05	66.1686	0.3424
Indiana	1.30e+04	3.02e+05	23.2644	2.8538
Iowa	5.65e+04	2.80e+06	49.6587	12.4210
Kansas	6.84e+04	3.02e+04	0.4420	15.0472
Kentucky	1.39e+04	1.66e+05	11.9861	3.0639
Louisiana	3307.2636	9826.1476	2.9711	0.7274
Maine	1.13e+04	9.44e+05	83.8732	2.4753
Maryland	2.15e+04	2.60e+05	12.1083	4.7307

In-class example: synth2 output

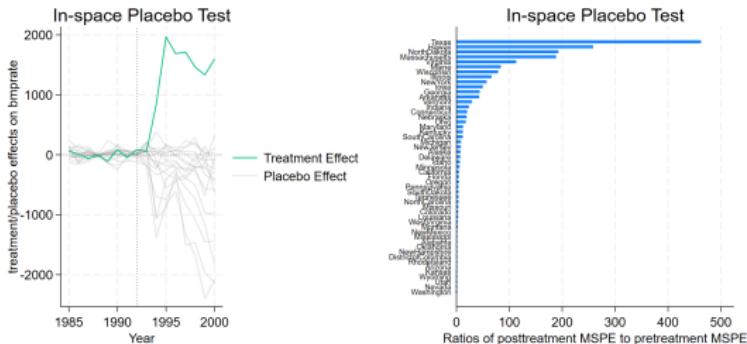


In-class example: “in-space” placebo test

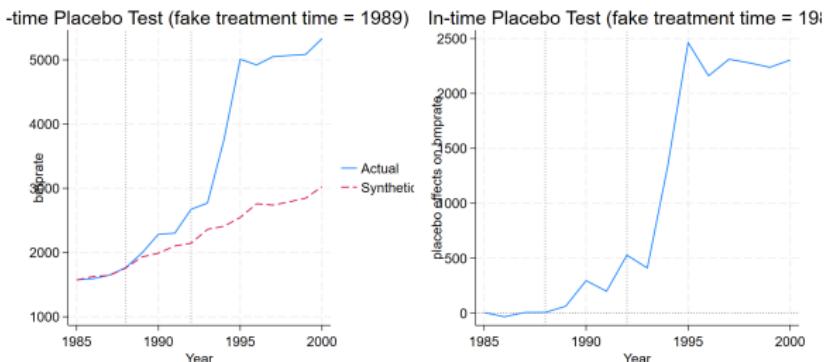


In-class example: “in-space” placebo test

Note: this case omits placebo states with a particularly bad pre-treatment fit (using cutoff suboption).

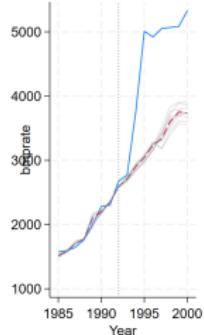


In-class example: “in-time” placebo test

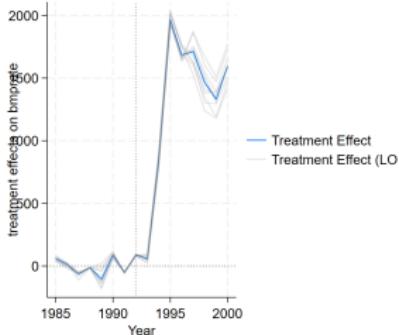


In-class example: leave-one-out robustness test

Leave-one-out Robustness Test



Leave-one-out Robustness Test



Extensions

The synthetic control literature is expanding rapidly and includes a number of useful extensions:

- Allowing for more than one treated unit, and multiple treatment periods (Cavallo et al., 2013)
- Bias-corrected synthetic control methods: approaches to eliminating any remaining pre-treatment differences between the treated unit and synthetic control (e.g., see Ben-Michael, Feller, and Rothstein (2021) and Shores et al. (2023)).