Synthetic controls method

Lukáš Lafférs

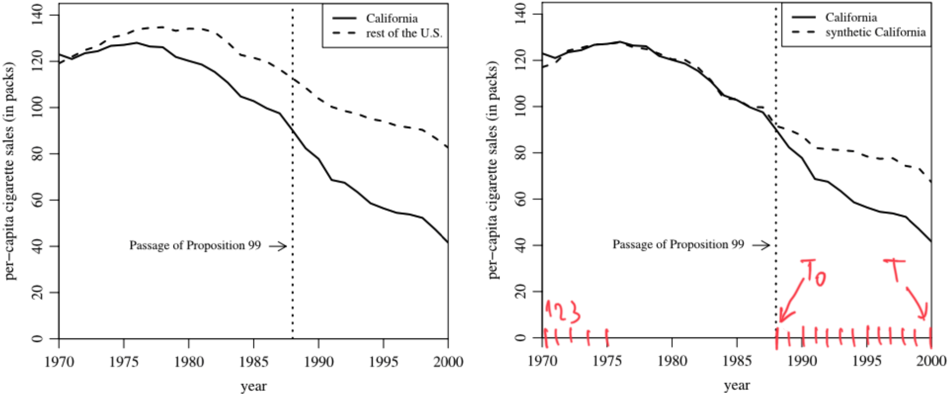
Matej Bel University, Dept. of Mathematics

MUNI Brno

~~3.12.2021~~ 7.1.2022

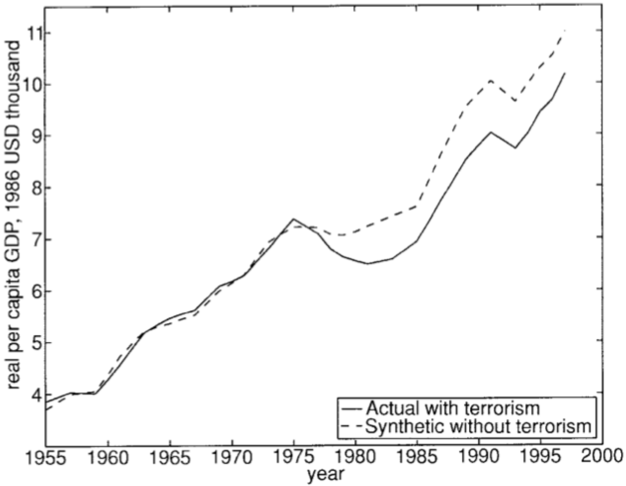
- In many situations the treatment happens on an aggregate level (city, state).
- We may not have a natural unit to use as a control
- We create it artificially (hence synthetic) by weighting other units so that the characteristics of the weighted unit resembles the one of the treated unit

Example: Tabacco control program and cigarettes sales



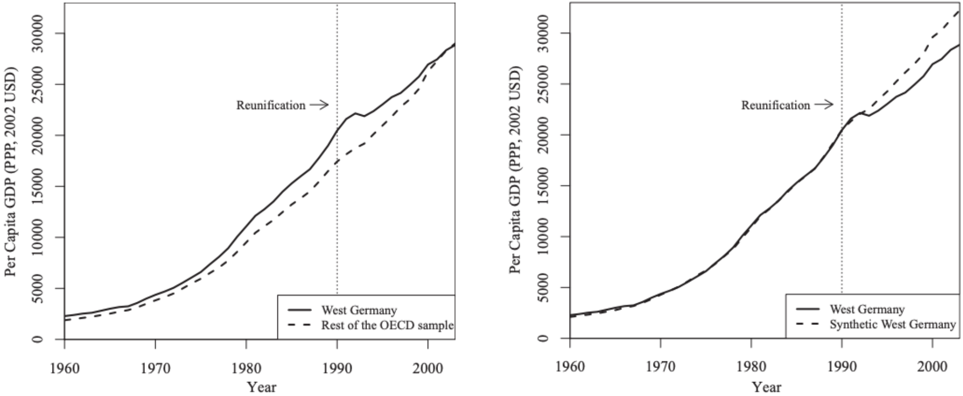
Abadie, Diamond and Hainmueller (2010)

Example: The economic cost of a conflict



Abadie and Gardezabal (2003)

Example: Reunification of Germany and Economic growth



Abadie, Diamond and Hainmueller (2015)

- time $1, 2, \ldots, T$
- $J + 1$ units, 1 is treated in $T_0 + 1, \ldots, T$
- Synthetic control is a weighted average of the J control units. $(w_2, \ldots, w_{J+1})$ with $w_j \geq 0, \sum_{j=2}^{J+1} w_j = 1$
- Weights w_j^* are chosen optimally to make the synthetic control similar to the control one in observed characteristics.
- Synthetic control estimator is

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

Choosing the weights

What does optimally mean?

We need some metric. Assume k variables $X_1, \ldots, X_k$. E.g. we can choose weighted Euclidean metric.

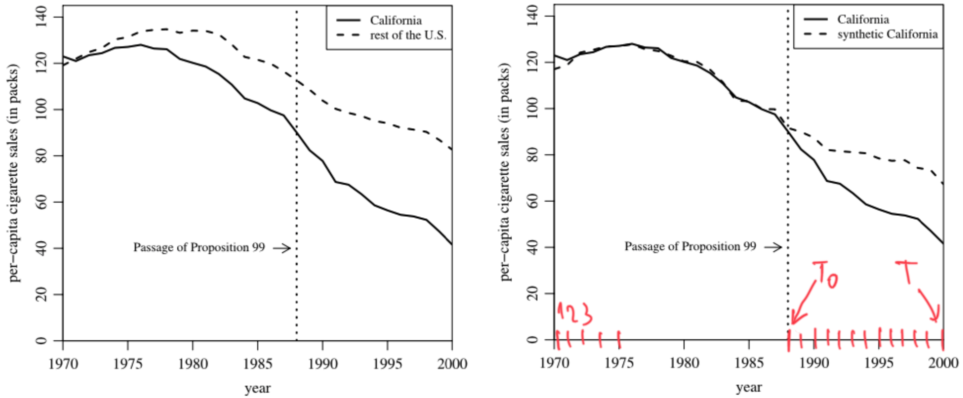
Pre-intervention outcomes are also included in the set of predictors!

Larger weights on more important predictors.

$$\arg\min_{w} \sum_{h=1}^{k} v_h \cdot \left(X_{h1} - \sum_{j=2}^{J+1} w_h \cdot X_{hj} \right)^2$$

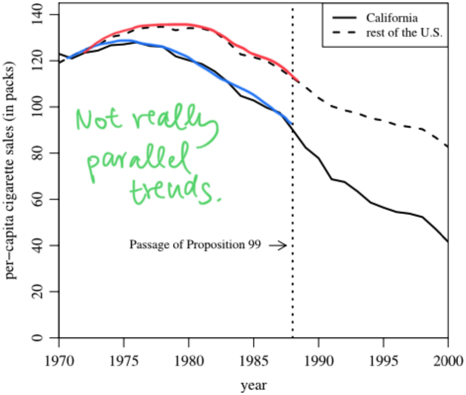
Assuming a linear factor model: If you manage to match controls and outcome in the pre-treatment periods ($T = 1, ..., T_0$) then you can bound the bias of the synthetic control method (Abadie, Diamond, and Hainmueller 2010).

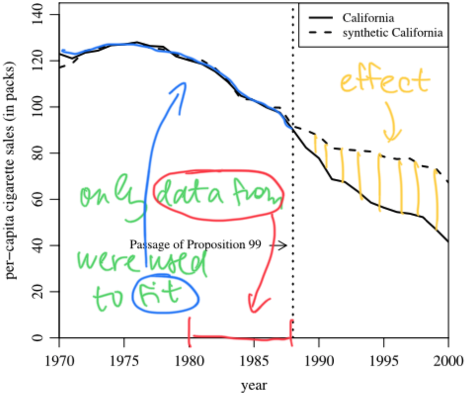
Example: Tabacco again



Abadie, Diamond and Hainmueller (2010)

Example: Tabacco again





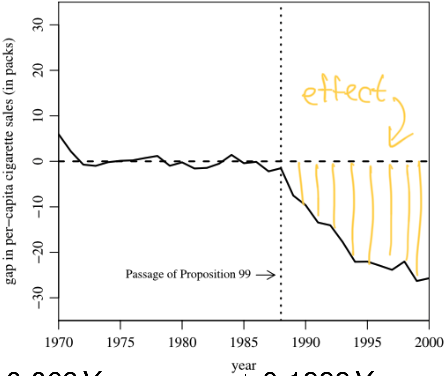
Abadie, Diamond and Hainmueller (2010)

Weights

Table 2. State weights in the synthetic California

State	Weight	State	Weight
Alabama	0	Montana	0.199
Alaska	–	Nebraska	0
Arizona	–	Nevada	0.234
Arkansas	0	New Hampshire	0
Colorado	0.164	New Jersey	–
Connecticut	0.069	New Mexico	0
Delaware	0	New York	–
District of Columbia	–	North Carolina	0
Florida	–	North Dakota	0
Georgia	0	Ohio	0
Hawaii	–	Oklahoma	0
Idaho	0	Oregon	–
Illinois	0	Pennsylvania	0
Indiana	0	Rhode Island	0
Iowa	0	South Carolina	0
Kansas	0	South Dakota	0
Kentucky	0	Tennessee	0
Louisiana	0	Texas	0
Maine	0	Utah	0.334
Maryland	–	Vermont	0
Massachusetts	–	Virginia	0
Michigan	–	Washington	–
Minnesota	0	West Virginia	0
Mississippi	0	Wisconsin	0
Missouri	0	Wyoming	0

sparse (many zeros)



$$
\begin{aligned}
Y_{synth,t} &= 0.164\,Y_{Colorado,t} + 0.069\,Y_{Connecticut,t} + 0.1999\,Y_{Montana,t} + \\
&\quad 0.234\,Y_{Nevada,t} + 0.334\,Y_{Utah,t}
\end{aligned}
$$

$$
\underbrace{\hat{\tau}_{California,t}}_{\text{effect}} = \underbrace{Y_{California,t}}_{\text{real outcome}} - \underbrace{Y_{synth,t}}_{\text{synthetic control}}
$$

Balance

Table 1. Cigarette sales predictor means

Variables	California Real	California Synthetic	Average 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

nice!

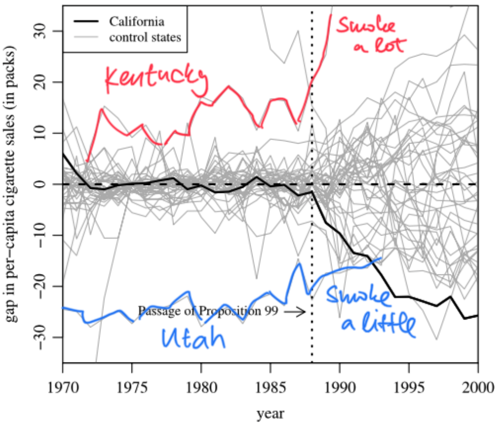
Abadie, Diamond and Hainmueller (2010)

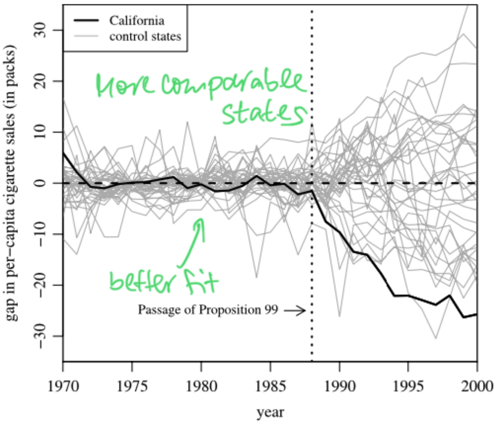
Inference

Use permutation method.

- Consider every control as a ”fake” treatment and estimate placebo effect
- Compare the effect for treated unit with those placebo effects
- Effect for the treated should be much larger than the placebo units
- But the pre-treatment fits may be different for different control units
- Abadie et al. (2010) suggests to look a the distribution of ratio of post vs pre-treatment fit
- Yes, we look at the whole distribution, not only p-values.

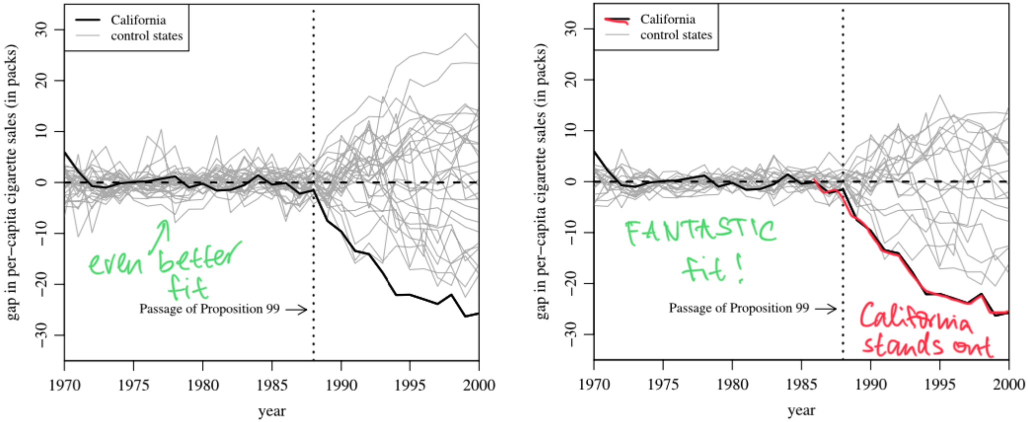
Placebos





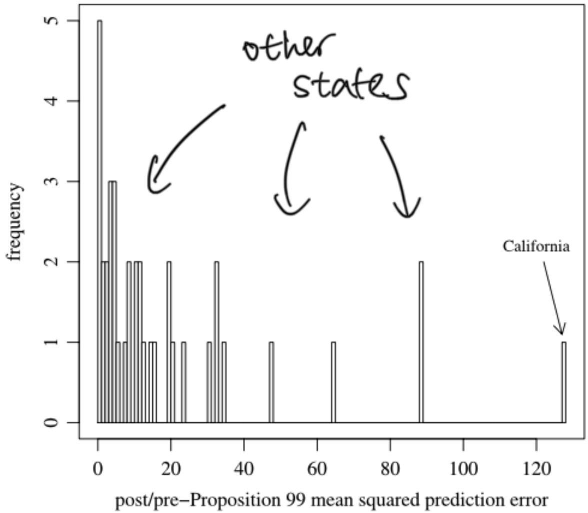
Abadie, Diamond and Hainmueller (2010)

Placebos



Abadie, Diamond and Hainmueller (2010)

Inference



Abadie, Diamond and Hainmueller (2010)

- If the fit is poor in the pre-intervention period. Do not do SCM, do something else.
- Small T_0 and large $J \rightarrow$ risk of overfitting
- Homogenise your pool of potential controls. Make them similar to the control unit.
- Again make comparison more plausible.

But why not regression instead?

Predictors X_0 (with intercept) are used to predict $y_{0,t}$ (post intervention outcomes for J control units at time $t \in T_0 + 1, \ldots, T$):

$$
\begin{aligned}
\hat{\beta}_{OLS,t} &= (X_0^T X_0)^{-1} X_0^T y_{0,t} \\
\underbrace{X_1}_{1\times K} \underbrace{\hat{\beta}_{OLS,t}}_{K\times 1} &= \underbrace{X_1 (X_0^T X_0)^{-1} X_0^T}_{w^T \equiv \text{ OLS weights}} y_{0,t} = \underbrace{w^T}_{1\times J} \underbrace{y_{0,t}}_{J\times 1}
\end{aligned}
$$

Let us denote $Y_0 = \begin{bmatrix} y_{0,T_0+1} & y_{0,T_0+2} & \cdots & y_{0,T} \end{bmatrix}$ which is $J \times (T - T_0)$ matrix.

$$
\begin{aligned}
\underbrace{\hat{B}_{OLS}}_{K\times(T-T_0)} &= (\underbrace{X_0^T}_{K\times J} \underbrace{X_0}_{J\times K})^{-1} \underbrace{X_0^T}_{K\times J} \underbrace{Y_0}_{J\times(T-T_0)} \\
\underbrace{X_1}_{1\times K} \underbrace{\hat{B}_{OLS}}_{K\times(T-T_0)} &= \underbrace{X_1 (X_0^T X_0)^{-1} X_0^T}_{w^T \equiv \text{ OLS weights}} Y_0 = \underbrace{w^T}_{1\times J} \underbrace{Y_0}_{J\times(T-T_0)}
\end{aligned}
$$

But why not regression instead?

TABLE 2
SYNTHETIC CONTROL WEIGHTS FOR WEST GERMANY

Australia	—
Austria	0.42
Belgium	—
Denmark	—
France	—
Greece	—
Italy	—
Japan	0.16
Netherlands	0.09
New Zealand	—
Norway	—
Portugal	—
Spain	—
Switzerland	0.11
United Kingdom	—
United States	0.22

TABLE 3
REGRESSION WEIGHTS FOR WEST GERMANY

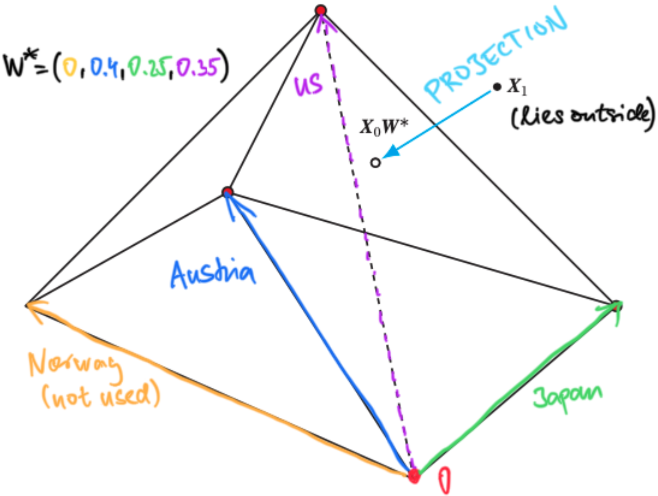
Australia	0.12
Austria	0.26
Belgium	0.00
Denmark	0.08
France	0.04
Greece	−0.09
Italy	−0.05
Japan	0.19
Netherlands	0.14
New Zealand	0.12
Norway	0.04
Portugal	−0.08
Spain	−0.01
Switzerland	0.05
United Kingdom	0.06
United States	0.13

weights negative (?)

Abadie (2021)

- From OLS we have also weights (!)
- May be negative $\rightarrow$ difficult to interpret
- OLS weights are not sparse
- Sparsity is nice for interpretation

Sparsity?



Abadie (2021)

Induce sparsity (penalized estimator)

We may induce the sparsity, so penalize for large differences.

$$\arg\min_{w} \underbrace{\left(\sum_{h=1}^{k} v_h \cdot \left(X_{h1} - \sum_{j=2}^{J+1} w_h \cdot X_{hj} \right)^2 \right)^{\frac{1}{2}}}_{\text{Regular SCM}} + \underbrace{\lambda \left(\sum_{j=2}^{J+1} w_h \sum_{h=1}^{k} v_h \cdot (X_{h1} - X_{hj})^2 \right)^{\frac{1}{2}}}_{\text{Penalty for non-sparse solution}}$$

We are in between the two extreme cases:

- $\lambda \to 0$ - synthetic control method
- $\lambda \to \infty$ - nearest neighbor matching

Alberto Abadie on DAGs

"Synthetic controls,... like in any other method for causal inference, what you won't be able to do is to whisper a question in a microphone to a computer and DAG will produce the answer for you. You have to make design decisions about what is a good comparison and what is not. And that's the case here too."

(Abadie in `https://www.youtube.com/watch?v=nKzNp-qpE-I` (from 59:50))

Advantages

- No extrapolation is made
- The weights make it transparent
- We know exactly how much each control unit contributes
- Weights are non-negative (unlike for OLS)
- You can fix the weights before the change has occurred.
- Thus you avoid specification fishing.
- You don't need many units, but the right units
- You are relatively close to the data $\rightarrow$ the method is simple

We keep getting back to the most important question:

What do we need to do in order to have a meaningful comparison?

What do many of these methods (RDD, DiD, SCM) have in common??

[dramatic pause]

They are very visual.

Professional graphics sells. Make sure to produce beautiful graphs. (See the works of Jonathan Schwabish on how to make great visualizations).

- Schwabish, Jonathan A. ”An economist’s guide to visualizing data.” Journal of Economic Perspectives 28.1 (2014): 209-34.
- Schwabish, Jonathan. Better presentations. Columbia University Press, 2016.
- Schwabish, Jonathan. Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks. Columbia University Press, 2021.

Synthetic controls and experimentation

- What is the impact of a new policy?
- We can only experiment on larger units (say cities).
- We choose some units (cities) and weight them to construct synthetic treatment unit, that resembles the population of interest.
- Construct synthetic control unit for this synthetic treatment unit
- And compare them. Yes, that's it.
- This has been used in the industry for a longer time.
- Abadie and Zhao (2021) worked out the math.

- SCM is new
- It is very popular and constantly getting more traction
- Much will be done in the next few years
- It became a standard in econometrics toolbox

Thank you for your attention!

References

- Original paper: Abadie, Alberto, and Javier Gardeazabal. "The economic costs of conflict: A case study of the Basque Country." American economic review 93.1 (2003): 113-132.
- Paper where theory is worked out: Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." Journal of the American statistical Association 105.490 (2010): 493-505.
- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. "Comparative politics and the synthetic control method." American Journal of Political Science 59.2 (2015): 495-510.
- Recent review article: Abadie, Alberto. "Using synthetic controls: Feasibility, data requirements, and methodological aspects." Journal of Economic Literature 59.2 (2021): 391-425.
- Instructive video by inventor of SCM himself `https://www.youtube.com/watch?v=nKzNp-qpE-I`
- Similar, slightly longer video also by Abadie at 2021 NBER Summer Institute lecture series `https://www.youtube.com/watch?v=T2p9Wg650bY`
- Abadie, Alberto, and Jinglong Zhao. "Synthetic controls for experimental design." arXiv preprint arXiv:2108.02196 (2021).
- Chapter 10 in S.Cunningham's book: `https://mixtape.scunning.com/synthetic-control.html`
- It is not often that WSJ writes about econometric methods: `https://www.wsj.com/articles/how-an-analysis-of-basque-terrorism-helps-economists-understand-brexit-1541587068`