合成控制法
Synthetic Control Method

陆嘉炫
Jiaxuan Lu

Jiaxuan.Lu@chicagobooth.edu

Booth School of Business
University of Chicago

August 20, 2021

Roadmap

- A Practitioner's Guide to the SCM
  - Introduction & Intuitions
  - Model
  - Statistical Tests
  - Stata Commands
- A Summary of the Econometrics Literature of the SCM
  - Challenges & Research Frontiers
  - SCM and Other Methods
  - Variants of the SCM
  - Inference
  - Multiple Treated Units
- A Brief Introduction of My Research
- Q&A

Introduction

- The synthetic control method (SCM) was originally proposed by Abadie and Gardeazabal (2003) and Abadie *et. al.* (2010).
- The SCM is "arguably the most important innovation in the policy evaluation literature in the last 15 years" (Athey and Imbens, 2017).
- Examples of Economic Applications
  - Abadie *et. al.* (2010) for cigarette sales
  - Bohn *et. al.* (2014) for immigration policies
  - Pinotti (2015) for organized crime
  - Acemoglu *et. al.* (2016) for corporate political connections
- Examples of Applications in Other Social Science Disciplines
  - Abadie *et. al.* (2015) and Heersink *et. al.* (2017) in political science
  - Pieters *et. al.* (2017) in consumer research
- Media Coverage: *The Washington Post*; *The Wall Street Journal*

Intuitions

- There are two periods.
- The 1st period is pre-treatment, and the 2nd period is post-treatment.
- There are $J + 1$ regions.
- Treatment X affects region i's outcome Y in the second period.
- Other J regions' values of Y are very different from region i. We can't simply compare region i with any of these regions alone.
- For the first period, we find a weighted sum of Y across these J regions such that this weighted sum is approximately equal to region i's value of Y. (It could be more complicated than taking the sum as long as it best resembles the characteristics of region i.)

- We use these weights to calculate the weighted sum of the values of Y across the J regions in the second period.
- This weighted sum is the estimated counterfactual of region i's value of Y in the absence of the treatment X.
- Treatment X's (Treatment-on-the-Treated, i.e. **ToT**) effect on outcome Y is the observed value of region i's value of Y in the second period minus its counterfactual value in the second period.

- Treated Unit: GREEN
- Control Units: BLUE & YELLOW & RED & PINK
- Counterfactual of the Treated Unit: 0.5×BLUE+0.5×YELLOW

When Should We Use the Synthetic Control Method?

- The treatment only affects **ONE** or a few aggregate cross-sectional units (i.e. cities, regions, *etc.*).
- The **parallel trends assumption** does **NOT** hold true. (In other words, each cross-sectional unit in the donor pool (aka control group) alone is very different from the treated unit.)

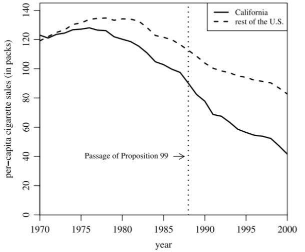


Figure 1. Trends in per-capita cigarette sales: California vs. the rest of the United States.

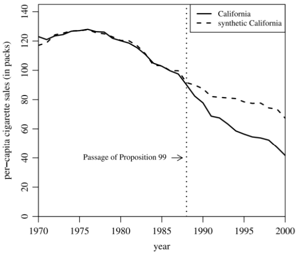


Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.

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

$$Y_{i,t} = Y_{i,t}^N + D_{i,t}$$
$$Y_{i,t}^N = \boldsymbol{\theta}_t \boldsymbol{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \varepsilon_{i,t}$$

- **Treated Unit**: $i = 0$
- **Donor Pool** (aka control group): $i = 1, 2, \cdots, J$
- **Pre-Treatment Period**: $t = 1, 2, \cdots, T_0$
- **Post-Treatment Period**: $t = T_0 + 1, T_0 + 2, \cdots, T$
- $D_{i,t}$: Treatment Effect
- $D_{i,t} > 0$ if $i = 0$ AND $t > T_0$; $D_{i,t} = 0$ if $i \neq 0$ OR $t \leq T_0$
- $Y_{i,t}$: **Observed** Outcome Variable
- $Y_{i,t}^N$: **Counterfactual** Outcome Variable
- $\varepsilon_{i,t}$: Idiosyncratic Error Term

Model

$$Y_{i,t} = Y_{i,t}^N + D_{i,t}$$
$$Y_{i,t}^N = \boldsymbol{\theta}_t \boldsymbol{Z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \varepsilon_{i,t}$$

- $\boldsymbol{\theta}_t$: $1 \times F$ Vector of **Observed Common Factors**
  - exchange rate affecting all regions
  - minimum wage affecting all workers
- $\boldsymbol{Z}_i$: $F \times 1$ Vector of **Observed Factor Loadings**
  - a region's export
  - a worker's educational attainment
- $\boldsymbol{\lambda}_t$: $1 \times R$ Vector of **Unobserved Common Factors**
  - financial crisis affecting all regions
  - price of an unmeasured skill affecting all workers
- $\boldsymbol{\mu}_i$: $R \times 1$ Vector of **Unobserved Factor Loadings**
  - a region's willingness to trade
  - a worker's ability

Model

- The SCM is to find the weight w_i^* for any $i \in \{1, 2, \cdots, J\}$ such that $Y_{0,t} \approx \sum_{j=1}^{J} w_j^* Y_{j,t}$ for $t \leq T_0$ and $\boldsymbol{Z}_0 \approx \sum_{j=1}^{J} w_j^* \boldsymbol{Z}_j$.
- Let $\boldsymbol{X}_j = [\{Y_{j,t}\}_{t=1}^{T_0}, \boldsymbol{Z}_j']'$. Abadie *et. al.* (2010) proposes the constrained optimization problem as follows:

$$\{\widehat{w}_1, \widehat{w}_2, \cdots, \widehat{w}_J\} = \underset{w_1, w_2, \cdots, w_J}{\arg\min} ||\boldsymbol{X}_0 - \sum_{j=1}^{J} w_j \boldsymbol{X}_j||$$

$$\text{subject to } w_j \geq 0 \text{ for any } j \in \{1, 2, \cdots, J\} \text{ and } \sum_{j=1}^{J} w_j = 1$$

- The **counterfactual** of the treated unit is $Y_{0,t}^N = \sum_{j=1}^{J} w_j^* Y_{j,t}$.
- The **pointwise estimator** is $\widehat{D}_{0,t} = Y_{0,t} - \widehat{Y}_{0,t}^N = Y_{0,t} - \sum_{j=1}^{J} \widehat{w}_j Y_{j,t}$.
- The **average treatment effect estimator** is $(T - T_0)^{-1} \sum_{t=T_0+1}^{T} \widehat{D}_{0,t}$.

Model

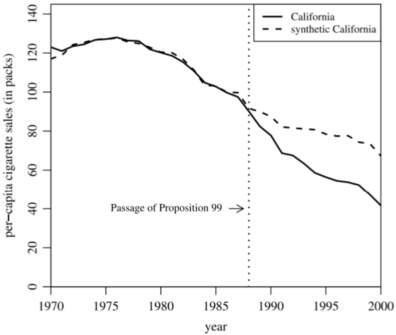


Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.

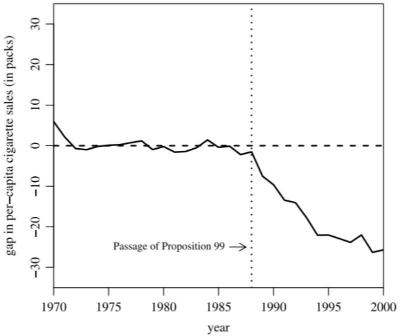


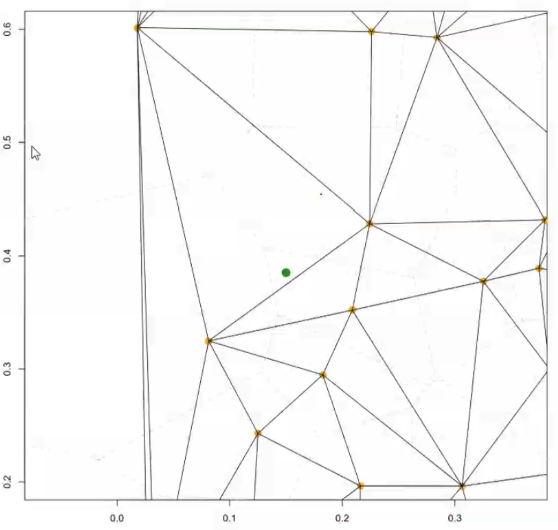
Figure 3. Per-capita cigarette sales gap between California and synthetic California.

Figures 2 and 3 from Abadie *et. al.* (2010)

Model

- **Assumptions and Requirements**
  - Comparable Donor Pool
  - No Interference: $Y_{i,t}$ and $Y_{j,\tau}$ can't affect one another for any $i \neq j$ and for any t and τ
  - No Anticipation: $Y_{i,t}$ is not affected in the pre-treatment period.
  - Sufficient Pre-Treatment Information: $T_0 \to \infty$
  - Sufficient Post-Treatment Information
- **Convex Hull Condition**
  - Abadie *et. al.* (2010): The weights are non-negative and sum-to-one.
  - This is not necessary, and it depends on your empirical context.
  - Chernozhukov *et. al.* (2021): The weights can be negative as long as the sum of the absolute values of the weights is smaller than one.
  - However, **sparsity** of the weights is highly recommended to avoid **over-fitting**.
- $\boldsymbol{Z}_i$ is not required, and it is probably difficult to find the data for $\boldsymbol{Z}_i$ in practice (especially for high-frequency data of $Y_{i,t}$).

Convex Hull Condition



Placebo Test

- Pretend that the post-treatment period starts from an earlier date $T_0'+1$ (where $T_0' < T_0$).
- Implement the SCM and estimate the treatment effect for the period between $T_0'+1$ and T_0. The estimated treatment effect should be zero.

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

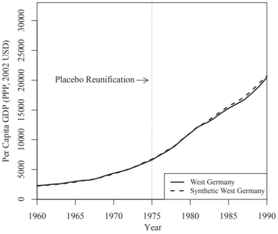


Figure 4 from Abadie *et. al.* (2015)

- GREEN= 0.5×BLUE+0.5×YELLOW
- The SCM estimation result suggests that GREEN receives a positive treatment effect, but is it possible that GREEN is unaffected while BLUE or YELLOW receives a negative treatment effect?
- We can iteratively reestimate the model to construct a synthetic treated unit omitting in each iteration one of the control units that receive a nonzero weight.

FIGURE 6 Leave-One-Out Distribution of the Synthetic Control for West Germany

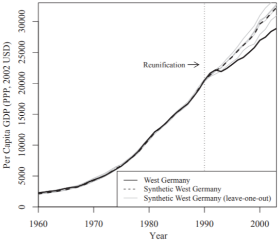


Figure 6 from Abadie *et. al.* (2015)

Inference

- We can obtain a **permutation distribution** by iteratively **reassigning** the treatment to a control unit in the donor pool and estimating placebo effects in each iteration.
- The estimated treatment effect on the treated unit is **statistically significant** if it is very "different" from the permutation distribution.

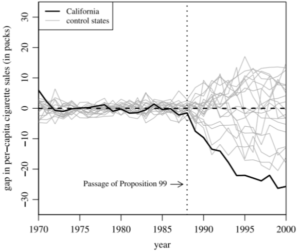
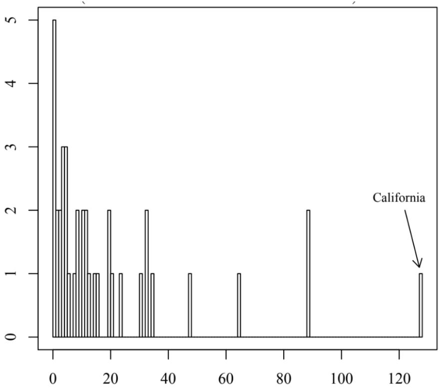


Figure 7. Per-capita cigarette sales gaps in California and placebo gaps in 19 control states (discards states with pre-Proposition 99 MSPE two times higher than California's).

Figure 7 from Abadie *et. al.* (2010)

Inference

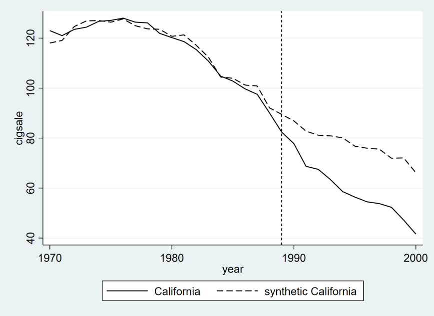
- MSPE: Mean Squared Prediction Error
- Prediction Error = Observed Value − Counterfactual Value
- In terms of the inference for the average treatment effect, one test statistic is the ratio of post-treatment MSPE to pre-treatment MSPE. This ratio should be larger for the treated unit in comparison to the control units.



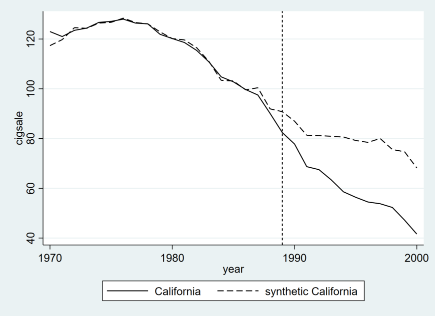
Stata Commands

- Research Question: What is the treatment effect of Proposition 99 in 1989 on California's cigarette sales?
- Abadie *et. al.* (2010)
- *synth cigsale lnincome(1980(1)1988) age15to24(1980(1)1988) \\\ retprice(1980(1)1988) beer(1984(1)1988) cigsale(1988) cigsale(1980) \\\ cigsale(1975), trunit(3) trperiod(1989) fig*
- My Preferred Approach
- *synth cigsale cigsale(1988) cigsale(1987) cigsale(1986) \\\ cigsale(1985) cigsale(1984) cigsale(1983) cigsale(1982) cigsale(1981) \\\ cigsale(1980) cigsale(1979) cigsale(1978) cigsale(1977) cigsale(1976) \\\ cigsale(1975) cigsale(1974) cigsale(1973) cigsale(1972) cigsale(1971) \\\ cigsale(1970), trunit(3) trperiod(1989) fig*

Stata Commands



Abadie *et. al.* (2010)



My Preferred Approach

- You could also use the *synth* package to create the tables as follows.

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

NOTE: All variables except lagged cigarette sales are averaged for the 1980–1988 period (beer consumption is averaged 1984–1988). GDP per capita is measured in 1997 dollars, retail prices are measured in cents, beer consumption is measured in gallons, and cigarette sales are measured in packs.

- You could also use the *synth* package to create the tables as follows.

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

Stata Commands

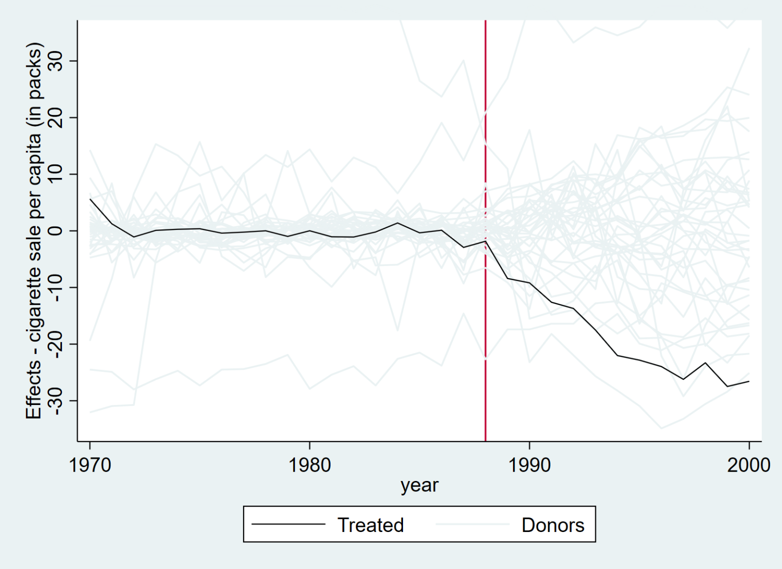
- synth_runner package (Galiani and Quistorff, 2017)
- This package is very friendly with pointwise inference.
- Take my preferred approach as an example.
- *synth_runner cigsale cigsale(1988) cigsale(1987) cigsale(1986) \\\ cigsale(1985) cigsale(1984) cigsale(1983) cigsale(1982) cigsale(1981) \\\ cigsale(1980) cigsale(1979) cigsale(1978) cigsale(1977) cigsale(1976) \\\ cigsale(1975) cigsale(1974) cigsale(1973) cigsale(1972) cigsale(1971) \\\ cigsale(1970), trunit(3) trperiod(1989)*

Stata Commands

	estimates	pvals	pvals_std
c1	-8.403796	.1052632	.0263158
c2	-9.178997	.2105263	.0789474
c3	-12.6086	.1315789	.1052632
c4	-13.7037	.1052632	.1052632
c5	-17.5067	.0789474	.0526316
c6	-22.0205	.0526316	.0526316
c7	-22.8257	.0526316	.0789474
c8	-23.9651	.0526316	.0526316
c9	-26.2214	.0789474	.0526316
c10	-23.2979	.0789474	.0526316
c11	-27.4804	.0526316	.0526316
c12	-26.5639	.0526316	.0526316

- c1: 1989, c2: 1990, c3: 1991, *etc.*
- estimates: pointwise SCM estimation result
- pvals: p-values calculated based on the permutation test (i.e. percentage of the control units with estimated placebo effect greater than the estimated effect received by the treated unit)
- pvals_std: p-values after standardization of the placebo estimates

Stata Commands



Resources for Applied Researchers

- 1. Chamberlain Seminar: Tutorial on Synthetic Control Methods (06/18/2021)

- 2. Abadie, A. (2021). "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature*, 59(2), 391-425.
- 3. Abadie, A., Diamond, A., & Hainmueller, J. (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490), 493-505.
- 4. Abadie, A., Diamond, A., & Hainmueller, J. (2015). "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*, 59(2), 495-510.
- 5. Galiani, S., & Quistorff, B. (2017). "The synth_runner Package: Utilities to Automate Synthetic Control Estimation Using synth." *Stata Journal*, 17(4), 834-849.
- 6. Hsiao, C., Ching, H. S., & Wan, S. K. (2012). "A Panel Data Approach for Program Evaluation: Measuring the Benefits of Political and Economic Integration of Hong Kong with Mainland China." *Journal of Applied Econometrics*, 27(5), 705-740.

Challenges & Research Frontiers

- *How shall we place the SCM in the statistics/econometrics literature?*
- *How shall we improve the pre-treatment fit for the purpose of creating better counterfacturals?*
- *How shall we make more formal statistical inference?*
- *How shall we use the SCM to evaluate the treatment effects when there are multiple treated units?*

Synthetic Control Method and Other Methods

- **SCM vs. Difference-in-Differences (DID)**
- The DID is a special case of the SCM where all cross-sectional units' weights are equal (Doudchenko and Imbens, 2016).
- The SCM could deal with factor models, but the DID can't do so (Bai, 2009; Gobillon and Magnac, 2016).
- The DID controls for the time fixed effects, but the original SCM does not do so. Arkhangelsky *et. al.* (2019) proposes to do so with the SCM by (for each cross-sectional unit) creating the counterfactuals for the post-treatment time periods by some linear combinations of pre-treatment time periods, i.e. $Y_{i,t} \approx \sum_{\tau=1}^{T_0} w_\tau Y_{i,\tau}$ for each post-treatment period t.
- If the number of treated units is small, then the SCM would be better than the traditional DID, but Conley and Taber (2011) develops an inference method that significantly improves the performance of DID when there is only a few treated units.

- **SCM vs. Interactive Fixed Effects (IFE)**
- Both methods could deal with factor models (Bai, 2009; Gobillon and Magnac, 2016).
- The SCM performs better when the number of treated units is very small. The IFE performs better when the number of treated units is relatively large.
- The IFE often requires the number of factors to be known. The SCM does not need this.
- Xu (2017) cleverly combines the SCM and the IFE so that his version of the SCM could evaluate the cases with multiple treated units.

- **SCM vs. Cointegration**
- Cointegration is similar to a special case of the SCM where only one control unit is selected for creating counterfactuals (Harvey and Thiele, 2021).
- Many time-series concepts should apply to the SCM.
- One needs to be cautious when applying the SCM to some nonstationary data (Bai *et. al.*, 2014; Masini and Medeiros, 2020; Masini and Medeiros, 2021).
- My working paper, i.e. Lu (2021), discusses nonstationarity and the pointwise inference of the SCM.

Variants of the Synthetic Control Method

The constrained optimization problem of the SCM could be generalized as follows.

$$\widehat{\boldsymbol{w}} = \underset{\boldsymbol{w} \in W}{\arg\min} ||\boldsymbol{X}_0 - \boldsymbol{X}\boldsymbol{w}||_p + \alpha||\boldsymbol{w}||_q$$

where $\boldsymbol{w}$ is the $J \times 1$ vector of weights, W is the domain for the weights, and $\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_j]$. $p = 2$ in most cases.

- Hsiao *et. al.* (2012): no observed factors $\boldsymbol{Z}_i$ in $\boldsymbol{X}_i$, $\alpha = 0$ (i.e. OLS)
- Doudchenko and Imbens (2016): no observed factors $\boldsymbol{Z}_i$ in $\boldsymbol{X}_i$, $||\boldsymbol{w}||_q$ is a weighted sum of $||\boldsymbol{w}||_1$ and $||\boldsymbol{w}||_2$ (i.e. elastic net)
- Li and Bell (2017): no observed factors $\boldsymbol{Z}_i$ in $\boldsymbol{X}_i$, $q = 1$ (i.e. LASSO)
- Li (2020): no observed factors $\boldsymbol{Z}_i$ in $\boldsymbol{X}_i$, $\boldsymbol{w} \geq 0$ (i.e. constrained OLS)
- Chernozhukov *et. al.* (2020; 2021): no observed factors $\boldsymbol{Z}_i$ in $\boldsymbol{X}_i$, $||\boldsymbol{w}||_1 \leq 1$ (i.e. constrained LASSO)

Variants of the Synthetic Control Method

- *How should we choose the value of the tuning parameter α?*
- **AICC** (Hsiao *et. al.*, 2012)
- Compute the AIC (corrected) score for each value of α. Choose the value of α with the smallest score.
- **Leave-One-Out** (Li and Bell, 2017)
- Given a value of α, predict the K-th pre-treatment period's outcome with all other pre-treatment periods' outcome data. Repeat this for $K = 1, 2, \cdots, T_0$. Choose the value of α with the smallest prediction error.
- **One-Step Ahead Forecast** (Kellogg *et. al.*, 2020)
- Given a value of α, use the outcome data of the first K pre-treatment periods to predict the $K+1$-th pre-treatment period's outcome. Repeat this for $K = T_0^0, T_0^0+1, \cdots, T_0$, where T_0^0 is a sufficiently large number (but still smaller than T_0). Choose the value of α with the smallest one-step ahead prediction error.

Pointwise Inference

- **Permutation Test** (Abadie *et. al.*, 2010)
- Plot the placebo estimates and compare them with the estimated treatment effect on the treated unit.
- **Rank Test** (Dube and Zipperer, 2015; Abadie and L'Hour, 2020)
- Use the percentile rank statistic in the permutation distribution.
- **OLS Prediction Intervals** (Hsiao *et. al.*, 2012)
- Use the OLS estimation results to create prediction intervals, but this is only limited to the panel data approach.

- **Permutation of the Prediction Errors** (Chernozhukov *et. al.*, 2021)
- If the treatment effect is zero, then the prediction errors should be i.i.d. and around zero across all periods. Permute the prediction errors across all periods within the treated unit for multiple iterations. The estimated treatment effect is significant in a post-treatment period if the magnitude of the actual prediction error for that period is larger than almost all iterations' permuted prediction errors for that period.

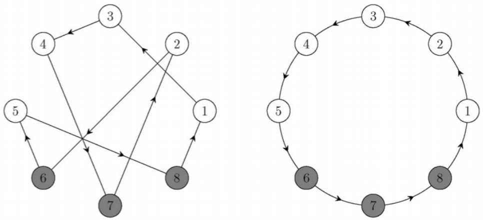


Figure 2. Graphical Illustration Permutations

Notes: The left figure gives an example of an iid permutation of $\{1,2,3,4,5,6,7,8\}$. The right figure gives an example of a moving block permutation of $\{1,2,3,4,5,6,7,8\}$. $T_0 = 5$, $T_* = 3$. Pretreatment periods are white; posttreatment periods are gray.

Figure 2 from Chernozhukov *et. al.* (2021)

Inference for the Average Treatment Effect

- The ATE estimator is $\widehat{D}_0 = (T - T_0)^{-1} \sum_{t=T_0+1}^{T} \widehat{D}_{0,t}$, and $E[\widehat{D}_0] = D_0$.
- **Ratio of Post-RMSPE to Pre-RMSPE** (Abadie *et. al.*, 2015; Firpo and Possebom, 2016)
- This method simply compares the treated unit's value of this ratio with other units' values of this ratio.
- **Asymptotic Distribution of the ATE Estimator** (Li and Bell, 2017; Li, 2020)

$$\sqrt{T - T_0}(\widehat{D}_0 - D_0) = -\sqrt{\frac{T - T_0}{T_0}} \frac{\sum_{t=T_0+1}^{T} \boldsymbol{X}_t}{T - T_0} \sqrt{T_0}(\widehat{\boldsymbol{w}} - \boldsymbol{w}) + \frac{1}{\sqrt{T - T_0}} \sum_{t=T_0+1}^{T} D_{0,t} - D_0 + \varepsilon_{0,t}$$

where both RHS terms converge to normal distributions based on some regularity conditions if unconstrained regression is used (e.g. Hsiao *et. al.*, 2012; Li and Bell, 2017). If constrained regression is used (e.g. Abadie *et. al.*, 2010; Li, 2020), then approximate the asymptotic distributions of the two RHS terms by the subsampling method.

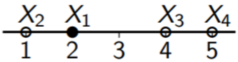
Multiple Treated Units

- **Generalized SCM** (Xu, 2017)
- Use the donor pool's data to estimate an interactive fixed effects (IFE) model. Use this IFE model to compute the factor loadings for each treated unit. Create counterfactuals based on these factor loadings and also the common factors estimated by the IFE model.
- Treated counterfactuals can be obtained in a single run.
- **Penalized SCM** (Abadie and L'Hour, 2020; Kellogg *et. al.*, 2020)

$$\widehat{\boldsymbol{w}} = \underset{\boldsymbol{w} \in W}{\arg\min} ||\boldsymbol{X}_0 - \boldsymbol{X}\boldsymbol{w}|| + \lambda \sum_{j=1}^{J} w_j ||\boldsymbol{X}_0 - \boldsymbol{X}_j||$$

- The second term is similar to a nearest neighbor matching estimator. It reduces the interpolation bias.
- Solution is guaranteed to be unique and sparse, and it is still computationally efficient.

Multiple Treated Units



- $X_1 = 2$ and $X_0 = [1\ 4\ 5]$.
- The (unpenalized) synthetic control has two sparse solutions: $W_1^* = (3/4,\ 0,\ 1/4)$ and $W_1^{**} = (2/3,\ 1/3,\ 0)$.
- W_1^{**} dominates W_1^* in terms of matching discrepancy. Infinite number of non-sparse solutions from convex combinations of these two.
- However, when $\lambda > 0$, the penalized synthetic control has a unique solution:

$$W_1^*(\lambda) = \begin{cases} (2+\lambda/2,\ 1-\lambda/2,\ 0)/3 & \text{if } 0 < \lambda \leq 2, \\ (1,\ 0,\ 0), & \text{if } \lambda > 2. \end{cases}$$

- As $\lambda \to 0$, $W_1^*(\lambda) \to W_1^{**}$, the pure synthetic control. The penalized synthetic control never uses the "bad" match X_4.

Page 39 of Alberto Abadie's Chamberlain Seminar Talk on June 18, 2021

Something More

- **Matrix Completion** (Athey *et. al.*, 2021)

$$\mathbf{Y}_{N\times T} = \begin{pmatrix} \checkmark & \checkmark & \checkmark & \checkmark & \dots & \checkmark & & \text{(never adopter)} \\ \checkmark & \checkmark & \checkmark & \checkmark & \dots & ? & & \text{(late adopter)} \\ \checkmark & \checkmark & ? & ? & \dots & ? & & \\ \checkmark & \checkmark & ? & ? & \dots & ? & & \text{(medium adopter)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & & \\ \checkmark & ? & ? & ? & \dots & ? & & \text{(early adopter)} \end{pmatrix}$$

- Use the nuclear norm to fill the question marks in the matrix.
- **Bias Correction** (Ben-Michael *et. al.*, 2021)
- $Y_{i,t}^N = m_{i,t} + u_{i,t}$, where $m_{i,t}$ is something observed (e.g. observed determinants of $Y_{i,t}^N$, similar to $\boldsymbol{\theta}_t \boldsymbol{Z}_i$).
- The bias-corrected SCM estimator is simply $\widehat{Y}_{0,t}^N = \sum_{j=1}^J w_j Y_j + (\widehat{m}_{0,t} - \sum_{j=1}^J w_j \widehat{m}_{j,t})$

Brief Introduction of My Research

- **Lu, J. (2021). "Synthetic Control Method, Stationarity and Pointwise Statistical Inference." Available at SSRN 3779281.**
- Any comments/suggestions/questions are welcome!
- What if?
  - The treatment effect is NOT long-lasting.
  - The treatment's exact beginning and ending dates are NOT known.
  - However, the beginning date is after a certain known date.
- If so, when using the synthetic control method, pointwise statistical inference needs to be made.

Brief Introduction of My Research

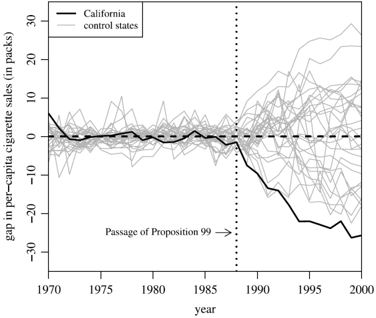


Figure 6 of Abadie *et. al.* (2010)
Outcome Variable: Cigarette Sales

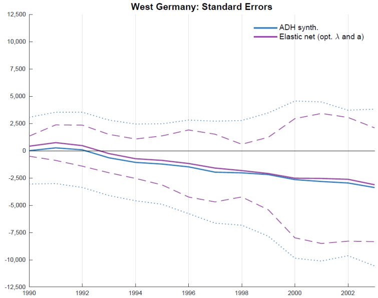


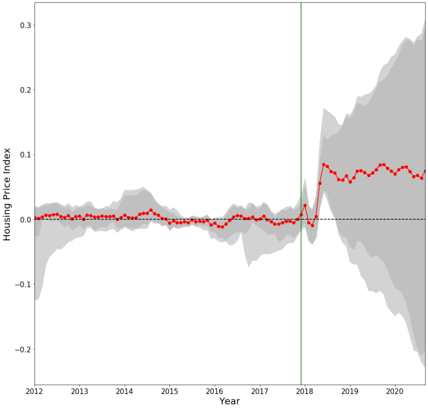
Figure 3 of Doudchenko and Imbens (2016)
Outcome Variable: GDP

- Why does the length of the confidence interval increase over time after the cutoff date?
- How should we make valid pointwise inference?

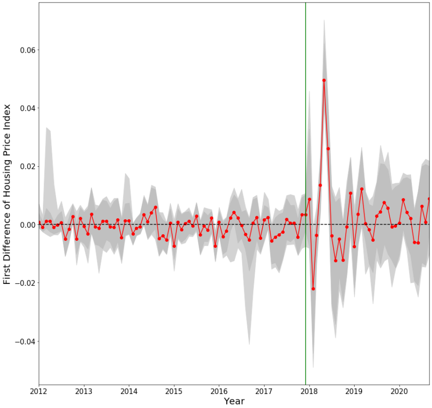
Brief Introduction of My Research

- The pointwise synthetic control estimator is still unbiased even if the outcome variable follows an autoregressive process.
- The variance of the pointwise estimator would be bounded if the process is stationary.
- The variance would be sensitive to the choice of the cutoff date and also increase over time unboundedly beginning from the cutoff date if the process is nonstationary.
- This nonstationarity should be removed by, for instance, first-differencing before using the synthetic control method.

Brief Introduction of My Research



(a) Before First-Differencing



(b) After First-Differencing

Outcome Variable: Housing Price Index

Closing Remarks

- The SCM is great, but it is not a panacea.
- Please use the SCM carefully and responsibly.
- Running a Stata code with the synth_runner command is easy, but it can't help me decide whether the use of the SCM in your research is a good or bad choice.
- It would be very helpful to understand the relationship between the SCM and other econometric methods.
- Think about why the SCM is better than other methods (such as DID) in your research.
- The SCM and its related methods deserve more theoretical explorations for econometricians.