Uncertainty Quantification in Synthetic Controls with Staggered Treatment Adoption

Matias D. Cattaneo, Princeton University

Yingjie Feng, Tsinghua University

Filippo Palomba, Princeton University

Rocio Titiunik, Princeton University

March 2023

Introduction

Synthetic control (SC) method.

- Introduced by Abadie and Gardeazabal (2003).
- Causal effect of a treatment on a single unit.
- Popular in comparative case studies.
  - German reunification, California Tobacco Control, Compulsory Voting, etc.
- Key idea: Predict $Y(0)$ of the treated by a linear combination of control units.
- Today: canonical single treatment unit case.
  - Paper/Software: multiple treatment units, staggered treatment adoption, more.

Our work focuses on uncertainty quantification for synthetic controls methods.

- Prediction Intervals for treatment effect on the treated.
- Two sources of uncertainty.
- Non-asymptotic coverage guarantees.
- Practical guidance.

Data Structure: Single Treatment Unit

Treatment: $D_{it} = 0, 1$ Outcome: $Y_{it} = Y_{it}(1)D_{it} + Y_{it}(0)(1 - D_{it})$

$$
\begin{array}{ccc}
 & \overbrace{\hphantom{\checkmark \quad \checkmark \quad \cdots \quad \checkmark}}^{\text{pre-treatment}} & \\
\text{unit} & \left(\begin{array}{cccc:ccc}
\checkmark & \checkmark & \cdots & \checkmark & \times & \cdots & \times \\
\hdashline
\checkmark & \checkmark & \cdots & \checkmark & \checkmark & \cdots & \checkmark \\
\vdots & \vdots & & & \vdots & & \vdots \\
\checkmark & \checkmark & \cdots & \checkmark & \checkmark & \cdots & \checkmark
\end{array}\right) & \begin{array}{l} \rightarrow \text{treated} \\ \\ \rightarrow \text{untreated} \\ \\ \end{array} \\
 & \text{time} &
\end{array}
$$

$Y(0) : \checkmark; \qquad Y(1) : \times$

Synthetic Control: Basics

$$Y_{it} = \begin{cases} Y_{it}(0) & \text{if } i = 2, \ldots N+1 \\ Y_{it}(0) & \text{if } i = 1 \text{ and } t \in \{1, 2, \ldots, T_0\} \\ Y_{it}(1) & \text{if } i = 1 \text{ and } t \in \{T_0 + 1, \ldots, T_0 + T_1\}. \end{cases}$$

Treatment effect on the treated (random!)

$$\tau_T = Y_{1T}(1) - Y_{1T}(0), \quad \text{for} \quad T > T_0$$

Find $\{w_i\}$

$$\sum_{i=2}^{N+1} w_i Y_{it}(0) \approx Y_{1t}(0), \quad \text{for} \quad t = 1, \cdots, T_0,$$

Hopefully,

$$\sum_{i=2}^{N+1} w_i Y_{iT}(0) \approx Y_{1T}(0), \quad \text{for} \quad T > T_0.$$

- Intuition: stable cross-sectional relation over time

Synthetic Control: Basics

$$Y_{it} = \begin{cases} Y_{it}(0) & \text{if } i = 2, \ldots N+1 \\ Y_{it}(0) & \text{if } i = 1 \text{ and } t \in \{1, 2, \ldots, T_0\} \\ Y_{it}(1) & \text{if } i = 1 \text{ and } t \in \{T_0 + 1, \ldots, T_0 + T_1\}\,. \end{cases}$$

Treatment effect on the treated (random!)

$$\tau_T = Y_{1T}(1) - Y_{1T}(0), \quad \text{for} \quad T > T_0$$

Find $\{w_i\}$

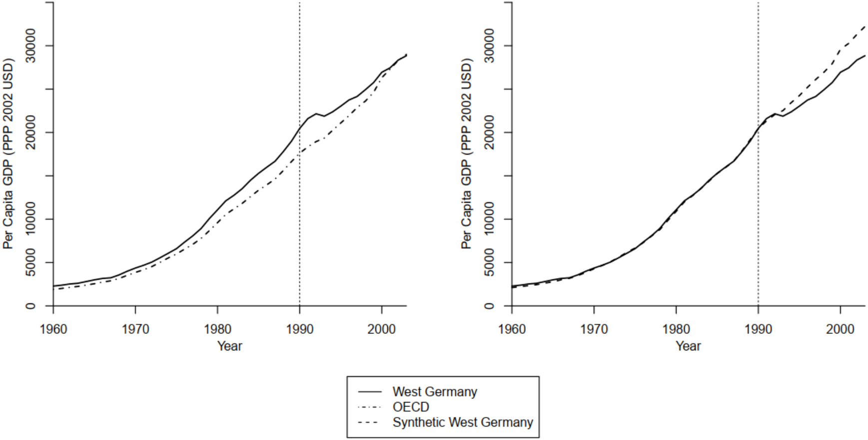
$$\sum_{i=2}^{N+1} w_i Y_{it}(0) \approx Y_{1t}(0), \quad \text{for} \quad t = 1, \cdots, T_0,$$

Canonical SC

$$\widehat{\boldsymbol{w}} = \underset{\boldsymbol{w} \in \mathcal{W}}{\arg\min} \sum_{t=1}^{T_0} \left(Y_{1t}(0) - \boldsymbol{x}_t' \boldsymbol{w} \right)^2,$$

- $\mathcal{W} = \{\boldsymbol{w} \in \mathbb{R}^N : \sum_{i=2}^{N+1} w_i = 1, w_i \geq 0, \ \forall i \geq 2\} \qquad \boldsymbol{x}_t = (Y_{2t}(0), \cdots, Y_{N+1,t}(0))'$

Example: 1990 German Reunification



Source: Abadie (2020)

Example: German Reunification

	West Germany (1)	Synthetic West Germany (2)	OECD Sample (3)
GDP per-capita	15808.9	15802.24	13669.4
Trade openness	56.8	56.9	59.8
Inflation rate	2.6	3.5	7.6
Industry share	34.5	34.5	34.0
Schooling	55.5	55.2	38.7
Investment rate	27.0	27.0	25.9

Note: First column reports $Y_1(0)$, second column reports $\boldsymbol{x}_t'\widehat{\boldsymbol{w}}$, and last column reports a simple average for the 16 OECD countries in the donor pool. GDP per capita, inflation rate, and trade openness are averages for 1981–1990. Industry share (of value added) is the average for 1981–1989. Schooling is the average for 1980 and 1985. Investment rate is averaged over 1980–1984.

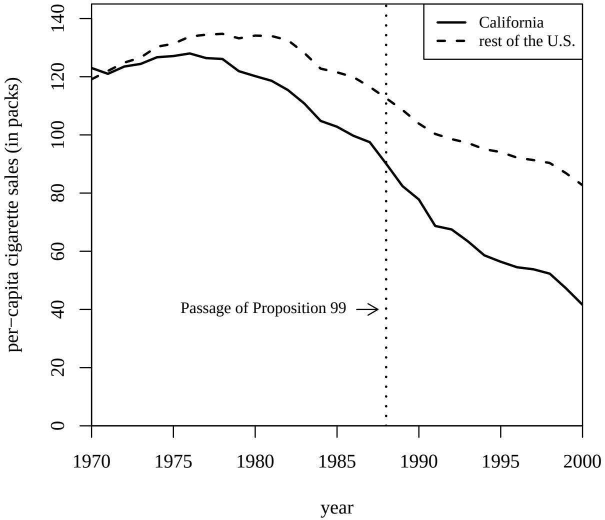
Example: German Reunification

country j	$[\widehat{\boldsymbol{w}}]_j$	country j	$[\widehat{\boldsymbol{w}}]_j$
Australia	0	Netherlands	0.10
Austria	0.42	New Zealand	0
Belgium	0	Norway	0
Denmark	0	Portugal	0
France	0	Spain	0
Greece	0	Switzerland	0.11
Italy	0	United Kingdom	0
Japan	0.16	United States	0.22

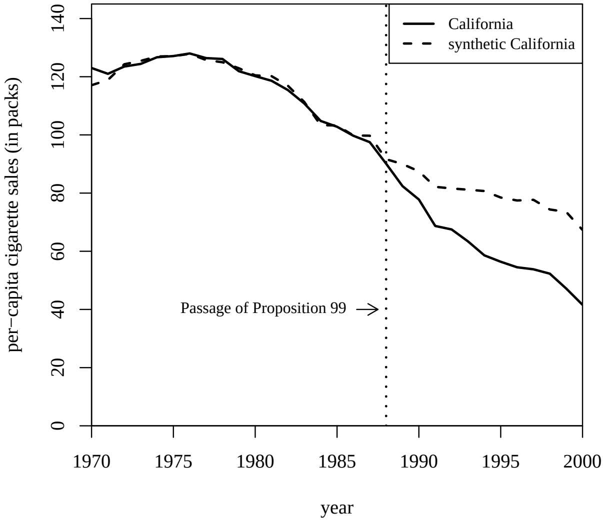
Uncertainty Quantification

- The quantity of interest: $\tau_T = Y_{1T}(1) - Y_{1T}(0)$, $T > T_0$, the effect of the intervention on the treated unit.
  - Random variable unless potential outcomes are assumed fixed (Neyman, Fisher)!
- How do we quantify uncertainty? Popular answer: Permutation-Based Inference.
  - Abadie et al. (2010) propose inference approach for the synthetic control framework based on permutation methods.
  - A permutation distribution can be obtained by iteratively reassigning the treatment to the units in the donor pool and estimating "placebo effects" in each iteration.
  - The effect of the treatment on the unit affected by the intervention is deemed to be significant when its magnitude is extreme relative to the permutation distribution.
  - The permutation distribution is more informative than mechanically looking at p-values alone.
  - Depending on the number of units in the donor pool, conventional significance levels may be unrealistic or impossible.
  - Often, one sided inference is most relevant.

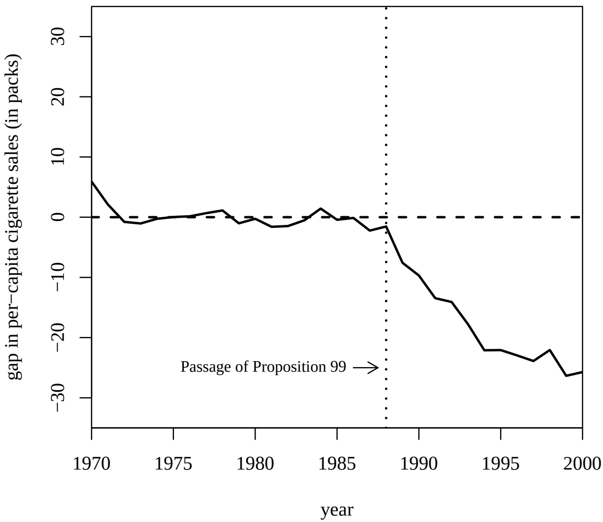
Example: California tobacco control program



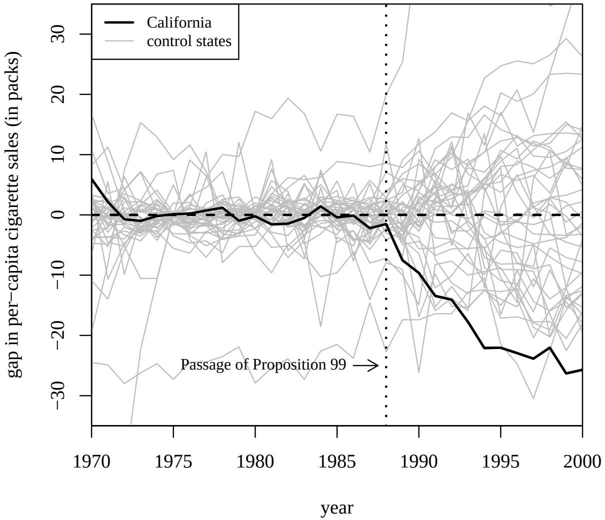
Example: California tobacco control program



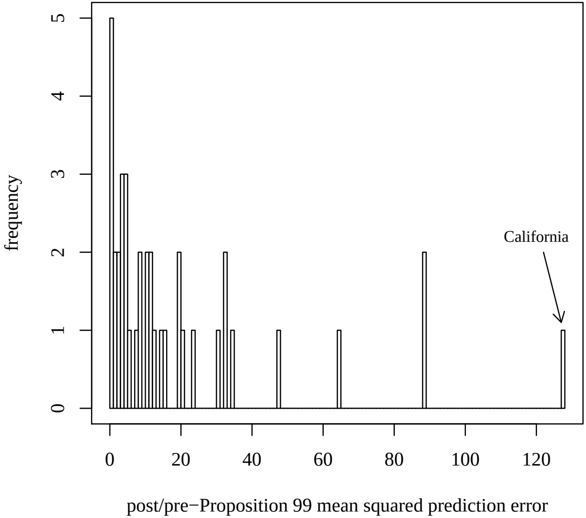
Example: California tobacco control program



Permutation-Based Inference



Example: California tobacco control program



Permutation-Based Inference in Synthetic Controls

- This mode of inference reduces to classical randomization inference (Fisher, 1935) when the intervention is randomly assigned, a rather improbable setting.
- More generally, this mode of inference evaluates significance relative to a benchmark distribution for the assignment process, one that is implemented directly in the data.
- But, it is not based on an a known assignment mechanism.
    - It is often difficult to articulate the nature of the assignment mechanism or even the specific nature of a placebo intervention (France reunifies with whom?).
    - Even if a plausible assignment mechanism exists, estimation of the assignment mechanism is often hopeless, because many studies feature a single or a small number of treated units.
    - Potential outcomes are more naturally conceptualized as random variables in most applications (interval vs. external validity).

Prediction Intervals

Treatment effect on the treated

$$\tau_T = Y_{1T}(1) - Y_{1T}(0), \qquad \widehat{\tau}_T = Y_{1T}(1) - \boldsymbol{x}_T'\widehat{\boldsymbol{w}}, \qquad T > T_0$$

Prediction interval (PI) for *random* τ_T

$$\mathbb{P}\Big\{\mathbb{P}\Big[\tau_T \in \mathcal{I} \,\Big|\, \mathscr{H}\Big] \geq 1-\alpha\Big\} \geq 1-\pi$$

- $\mathcal{I}$: prediction interval
- $\mathscr{H}$: conditioning σ-algebra
  - unconditional if $\mathscr{H}$ is trivial
  - This paper: $\{\boldsymbol{x}_t : 1 \leq t \leq T\}$
- $(1-\alpha)$: conditional coverage prob., e.g., 95%
- π: failure over $\mathscr{H}$

Prediction Intervals

Treatment effect on the treated

$$\tau_T = Y_{1T}(1) - Y_{1T}(0), \qquad \widehat{\tau}_T = Y_{1T}(1) - \boldsymbol{x}_T'\widehat{\boldsymbol{w}}, \qquad T > T_0$$

Prediction interval (PI) for *random* τ_T

non-asymptotic: $$\mathbb{P}\Big\{\mathbb{P}\Big[\tau_T \in \mathcal{I} \,\Big|\, \mathscr{H}\Big] \geq 1-\alpha\Big\} \geq 1-\pi$$

asymptotic: $$\mathbb{P}\Big[\tau_T \in \mathcal{C} \,\Big|\, \mathscr{H}\Big] \geq 1-\alpha - o_{\mathbb{P}}(1)$$

- Link the two
  - Consider $\pi = o(1)$ (as $T_0 \to \infty$)

Two Sources of Uncertainty

Treatment effect on the treated

$$\tau_T = Y_{1T}(1) - Y_{1T}(0), \qquad \widehat{\tau}_T = Y_{1T}(1) - \boldsymbol{x}_T'\widehat{\boldsymbol{w}}$$

For a pseudo-true value $\boldsymbol{w}_0$

$$Y_{1T}(0) = \boldsymbol{x}_T'\boldsymbol{w}_0 + u_T$$

A simple decomposition

$$\begin{aligned} \widehat{\tau}_T - \tau_T &= Y_{1T}(0) - \boldsymbol{x}_T'\widehat{\boldsymbol{w}} \\ &= (\boldsymbol{x}_T'\boldsymbol{w}_0 + u_T) - \boldsymbol{x}_T'\widehat{\boldsymbol{w}} \\ &= u_T - \boldsymbol{x}_T'(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0) \end{aligned}$$

- In-sample error: $\boldsymbol{x}_T'(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0)$
- Out-of-sample error: u_T
- Non-asymptotically, both are important

What is $\boldsymbol{w}_0$?

Definition of $\boldsymbol{w}_0$ relies on

- Constraint $\mathcal{W}$
- Assumptions on $\{Y_{1t}(0), \boldsymbol{x}_t\}$
- Conditioning set $\mathscr{H}$ ($\boldsymbol{w}_0$ may be random)

Examples

- $(Y_{1t}(0), \boldsymbol{x}_t)$ stationary

$$\boldsymbol{w}_0 = \underset{\boldsymbol{w} \in \mathcal{W}}{\arg\min}\, \mathbb{E}\Big[\frac{1}{T_0}\sum_{t=1}^{T_0}\Big(Y_{1t}(0) - \boldsymbol{x}_t'\boldsymbol{w}\Big)^2 \Big| \mathscr{H}\Big]$$

  - (Constrained) best linear prediction

- $Y_{1t}(0), \boldsymbol{x}_t \sim I(1)$ (integrated process)
  - Cointegration: $Y_{1t}(0) - \boldsymbol{x}_t'\boldsymbol{w}_0 \sim I(0)$
  - $(1, -\boldsymbol{w}_0)$: cointegrating vector (unique given $\boldsymbol{w}_0 \in \mathcal{W}$)

Prediction Intervals: Basic Construction

$$\widehat{\tau}_T - \tau_T = u_T - \boldsymbol{x}_T'(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0)$$

- **In-Sample Error**: with prob. $\geq 1 - \pi_1$ (over $\mathscr{H}$)

$$\mathbb{P}\Big[M_{1,\mathtt{L}} \leq \boldsymbol{x}_T'(\boldsymbol{w}_0 - \widehat{\boldsymbol{w}}) \leq M_{1,\mathtt{U}} \,\Big|\, \mathscr{H}\Big] \geq 1 - \alpha_1$$

- **Out-of-Sample Error**: with prob. $\geq 1 - \pi_2$ (over $\mathscr{H}$)

$$\mathbb{P}\Big[M_{2,\mathtt{L}} \leq u_T \leq M_{2,\mathtt{U}} \,\Big|\, \mathscr{H}\Big] \geq 1 - \alpha_2$$

Prediction Interval for τ_T: with prob. $\geq 1 - \pi_1 - \pi_2$ (over $\mathscr{H}$)

$$\mathbb{P}\Big[M_{1,\mathtt{L}} + M_{2,\mathtt{L}} \leq \widehat{\tau}_T - \tau_T \leq M_{1,\mathtt{U}} + M_{2,\mathtt{U}} \Big| \mathscr{H}\Big] \geq 1 - \alpha_1 - \alpha_2$$

- Conservative, but offer non-asymptotic probability guarantee

In-Sample Error: Stationary Case

$$\sqrt{T_0}(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0) = \operatorname*{arg\,min}_{\boldsymbol{\delta} \in \sqrt{T_0}(\mathcal{W} - \boldsymbol{w}_0)} \ell(\boldsymbol{\delta}), \quad \ell(\boldsymbol{\delta}) = \boldsymbol{\delta}' \underbrace{\Big(\frac{1}{T_0} \sum_{t=1}^{T_0} \boldsymbol{x}_t \boldsymbol{x}_t' \Big)}_{\widehat{\boldsymbol{Q}}} \boldsymbol{\delta} - 2 \underbrace{\Big(\frac{1}{\sqrt{T_0}} \sum_{t=1}^{T_0} \boldsymbol{x}_t' u_t \Big)}_{\widehat{\boldsymbol{\gamma}}} \boldsymbol{\delta}$$

- Assume $\mathcal{W}$ is convex
- Possibly misspecified: $\boldsymbol{\gamma} := \mathbb{E}[\widehat{\boldsymbol{\gamma}} | \mathscr{H}] \neq 0$
- By optimality of $\widehat{\boldsymbol{w}}$ and $\boldsymbol{w}_0$,

$$\sqrt{T_0}(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0) \in \Big\{ \boldsymbol{\delta} \in \Delta : \boldsymbol{\delta}' \widehat{\boldsymbol{Q}} \boldsymbol{\delta} - 2(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})' \boldsymbol{\delta} \leq 0 \Big\}, \quad \Delta = \sqrt{T_0}(\mathcal{W} - \boldsymbol{w}_0)$$

  - $\widehat{\boldsymbol{Q}}$ is fixed conditional on $\mathscr{H}$
  - Approximate $\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}$

In-Sample Error: Distributional Approximation

Goal: bound $\boldsymbol{x}_T'(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0)$, e.g.

$$\sqrt{T_0}\boldsymbol{x}_T'(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0) \leq \sup_{\boldsymbol{\delta} \in \mathcal{M}_{\widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}}} \boldsymbol{x}_T'\boldsymbol{\delta}$$

$$\mathcal{M}_{\widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}} = \left\{\boldsymbol{\delta} \in \Delta : \boldsymbol{\delta}'\widehat{\boldsymbol{Q}}\boldsymbol{\delta} - 2(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})'\boldsymbol{\delta} \leq 0\right\}$$

Reduce to distributional approximation of $\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}$: for any κ,

$$\sup_{\boldsymbol{\delta} \in \mathcal{M}_{\widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}}} \boldsymbol{x}_T'\boldsymbol{\delta} \leq \kappa \quad \Leftrightarrow \quad \widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \in \mathcal{A}_\kappa$$

- Normal approximation: Berry-Esseen bound

$$\mathbb{P}(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \in \mathcal{A}_\kappa) \approx \mathbb{P}(\boldsymbol{G} \in \mathcal{A}_\kappa), \quad \boldsymbol{G} \sim \mathsf{N}(0, \mathbb{V}[\widehat{\boldsymbol{\gamma}}|\mathscr{H}])$$

In-Sample Error: Plug-in Approximation

$$\mathbb{P}(\sqrt{T_0}\boldsymbol{x}_T'(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0) \leq \kappa) \geq \mathbb{P}(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma} \in \mathcal{A}_\kappa) \approx \mathbb{P}(\boldsymbol{G} \in \mathcal{A}_\kappa), \quad \boldsymbol{G} \sim \mathsf{N}(0, \mathbb{V}[\widehat{\boldsymbol{\gamma}}|\mathscr{H}])$$
$$\approx \mathbb{P}(\widehat{\boldsymbol{G}} \in \mathcal{A}_\kappa), \quad \widehat{\boldsymbol{G}} \sim \mathsf{N}(0, \widehat{\mathbb{V}}[\widehat{\boldsymbol{\gamma}}|\mathscr{H}])$$

Basic approximation strategy:

$$\sup\left\{\boldsymbol{x}_T'\boldsymbol{\delta} : \boldsymbol{\delta} \in \Delta, \;\; \boldsymbol{\delta}'\widehat{\boldsymbol{Q}}\boldsymbol{\delta} - 2(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})'\boldsymbol{\delta} \leq 0\right\}$$

$$\Downarrow$$

$$\sup\left\{\boldsymbol{x}_T'\boldsymbol{\delta} : \boldsymbol{\delta} \in \Delta, \;\; \boldsymbol{\delta}'\widehat{\boldsymbol{Q}}\boldsymbol{\delta} - 2\boldsymbol{G}'\boldsymbol{\delta} \leq 0\right\}$$

$$\Downarrow$$

$$\sup\left\{\boldsymbol{x}_T'\boldsymbol{\delta} : \boldsymbol{\delta} \in \Delta^\star, \boldsymbol{\delta}'\widehat{\boldsymbol{Q}}\boldsymbol{\delta} - 2\widehat{\boldsymbol{G}}'\boldsymbol{\delta} \leq 0\right\}$$

- $\Delta^\star$: locally equivalent to $\Delta = \sqrt{T_0}(\mathcal{W} - \boldsymbol{w}_0)$
- Thresholding

In-Sample Error: Non-Stationary Case

$$T_0(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0) = \underset{\boldsymbol{\delta} \in T_0(\mathcal{W} - \boldsymbol{w}_0)}{\arg\min} \ \ell(\boldsymbol{\delta}), \quad \ell(\boldsymbol{\delta}) = \boldsymbol{\delta}' \underbrace{\left(\frac{1}{T_0^2} \sum_{t=1}^{T_0} \boldsymbol{x}_t \boldsymbol{x}_t' \right)}_{\widehat{\boldsymbol{Q}}} \boldsymbol{\delta} - 2 \underbrace{\left(\frac{1}{T_0} \sum_{t=1}^{T_0} \boldsymbol{x}_t' u_t \right)}_{\widehat{\boldsymbol{\gamma}}} \boldsymbol{\delta}$$

Cointegrated system, e.g.,

$$Y_{1t}(0) = \boldsymbol{x}_t' \boldsymbol{w}_0 + u_t, \quad \boldsymbol{x}_t = \boldsymbol{x}_{t-1} + \boldsymbol{\epsilon}_t, \quad (u_t, \boldsymbol{\epsilon}_t) \sim \text{i.i.d } I(0)$$

Again, bound $T_0(\widehat{\boldsymbol{w}} - \boldsymbol{w}_0)$ by simulating

$$\sup \left\{ \boldsymbol{x}_T' \boldsymbol{\delta} : \boldsymbol{\delta} \in \Delta^\star, \ \boldsymbol{\delta}' \widehat{\boldsymbol{Q}} \boldsymbol{\delta} - 2 \widehat{\boldsymbol{G}}' \boldsymbol{\delta} \leq 0 \right\}, \quad \widehat{\boldsymbol{G}} \sim \mathsf{N}(0, \widehat{\mathbb{V}}[\widehat{\boldsymbol{\gamma}} | \mathscr{H}])$$

- $\widehat{\boldsymbol{Q}}$ fixed conditional on $\mathscr{H}$
- Probably $\mathbb{E}[u_t | \mathscr{H}] \neq 0$
- Non-stationarity affects the analysis of $\widehat{\boldsymbol{Q}}$ and $\mathbb{V}[\widehat{\boldsymbol{\gamma}} | \mathscr{H}]$

Out-of-Sample Error

Three approaches:

- Concentration inequalities, e.g. subgaussian

$$\mathbb{P}\Big(|u_T - \mathbb{E}[u_T|\mathscr{H}]| \geq \varpi_u|\mathscr{H}\Big) \leq 2\exp\Big(-\frac{\varpi_u^2}{2\sigma_{\mathscr{H}}^2}\Big)$$

- Location-scale model

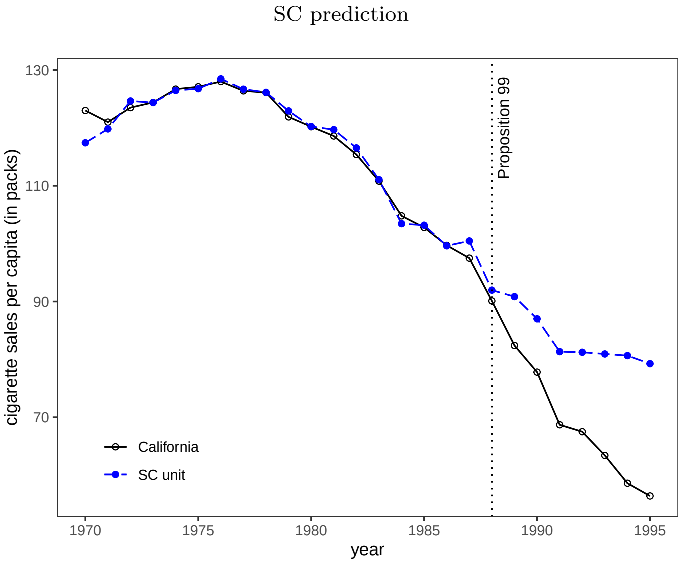
$$u_t = \mathbb{E}[u_t|\mathscr{H}] + \big(\mathbb{V}[u_t|\mathscr{H}]\big)^{1/2} e_t, \quad \{e_t\} \perp\!\!\!\perp \mathscr{H}$$

- Quantile regression: model conditional quantiles of u_t given $\mathscr{H}$

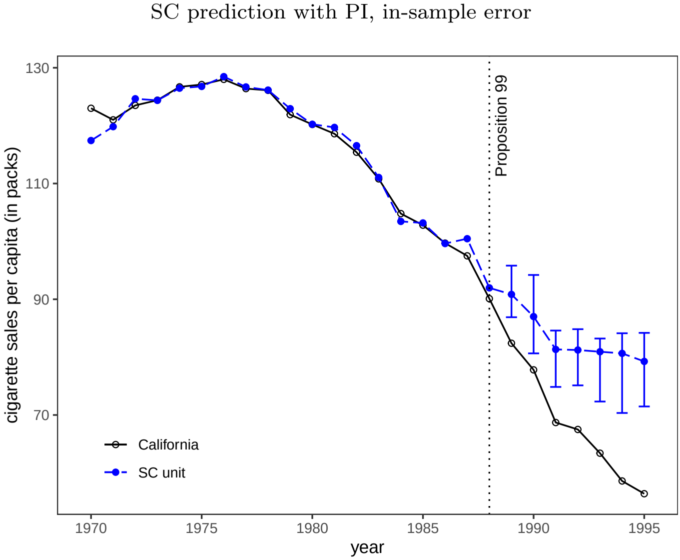
Sensitivity analysis

- To cancel out the effect, how large u_T needs to be?

Example: Proposition 99 (California)

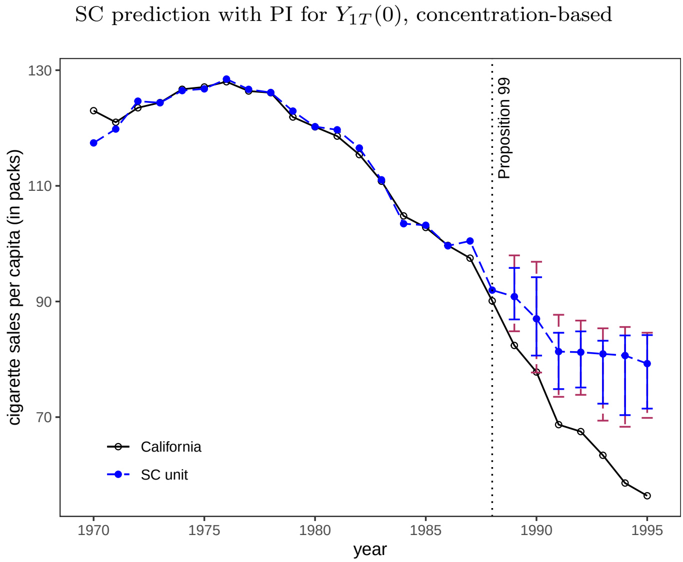


Example: Proposition 99 (California)



Note: PI has $\geq 95\%$ coverage.

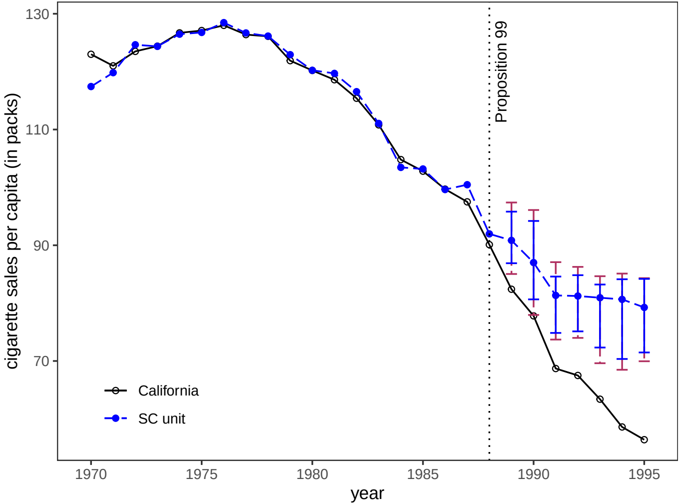
Example: Proposition 99 (California)



Note: PI has $\geq 90\%$ coverage.

Example: Proposition 99 (California)

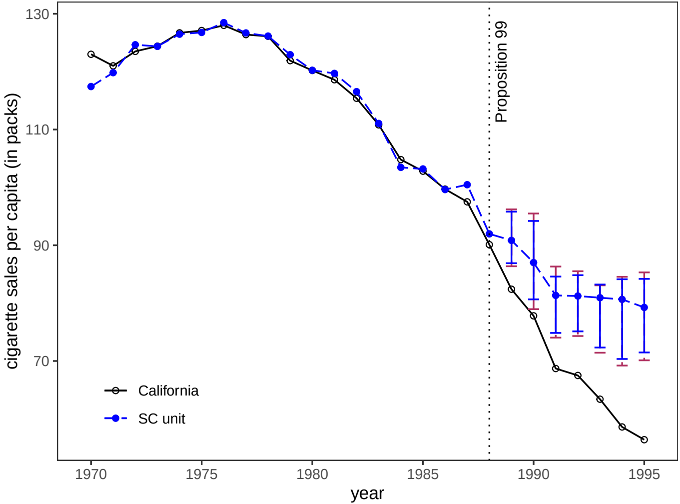
SC prediction with PI for $Y_{1T}(0)$, local-scale model



Note: PI has $\geq 90\%$ coverage.

Example: Proposition 99 (California)

SC prediction with PI for $Y_{1T}(0)$, quantile-based



Note: PI has $\geq 90\%$ coverage.

General Form

Some extensions

- Match on multiple predictors
- Add more covariates to regression
- Other constraints $\mathcal{W}$

Goal: find $\boldsymbol{w} = (w_2, \cdots, w_{N+1})'$ and $\boldsymbol{r}_j = (r_{1,j}, \cdots, r_{k,j})'$

$$\underbrace{\boldsymbol{A}_j}_{\substack{\text{treated} \\ T_0 \times 1}} \approx \underbrace{\boldsymbol{B}_j}_{\substack{\text{untreated} \\ T_0 \times N}} \boldsymbol{w} + \underbrace{\boldsymbol{C}_j}_{\substack{\text{covariates} \\ T_0 \times K}} \boldsymbol{r}_j, \qquad \text{for} \quad j = 1, \cdots, M,$$

$$(\widehat{\boldsymbol{w}}, \widehat{\boldsymbol{r}}) = \underset{\boldsymbol{w} \in \mathcal{W}, \boldsymbol{r} \in \mathcal{R}}{\arg\min} \, (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{w} - \boldsymbol{C}\boldsymbol{r})'(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{w} - \boldsymbol{C}\boldsymbol{r})$$

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \vdots \\ \boldsymbol{A}_M \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \\ \vdots \\ \boldsymbol{B}_M \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2 & \cdots & \boldsymbol{0} \\ \vdots & & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{C}_M \end{bmatrix}$$

Conclusion

- Prediction Intervals for general SC methods.
- Conditional validity.
- Non-asymptotic probability guarantees.
- Two sources of uncertainty.
- Joint prediction intervals for staggered treatment adoptions (e.g., multiple treatment units).
- Principled tuning parameter selection.

References

1. Cattaneo, Feng & Titiunik (2021): "Prediction Intervals for Synthetic Control Methods" *Journal of the American Statistical Association* 116(536): 1865-1880.

    - Pointwise prediction intervals for one single treatment unit.
    - Inference valid for locally linear constraints.

2. Cattaneo, Feng, Palomba & Titiunik (2022): "Uncertainty Quantification in Synthetic Controls with Staggered Treatment Adoption", arXiv:2210.05026.

    - Joint prediction intervals for staggered treatment adoptions (e.g., multiple treatment units treated at possibly different points in time).
    - Inference valid for locally quadratic constraints (e.g., ridge regression or L1-L2 regularization).
    - Principled tuning parameter selection and scalable, robust conic programming implementations.

3. Cattaneo, Feng, Palomba & Titiunik (2022): "`scpi`: Uncertainty Quantification for Synthetic Control Estimators", arXiv:2202.05984.

    - Software implementation: `https://nppackages.github.io/scpi/`