Advanced Synthetic Control Methods

Lee Kennedy-Shaffer, PhD

2025-06-10

Key Ideas

Key Ideas

- Allow for staggered adoption

Key Ideas

- Allow for staggered adoption
- Reduce interpolation

Key Ideas

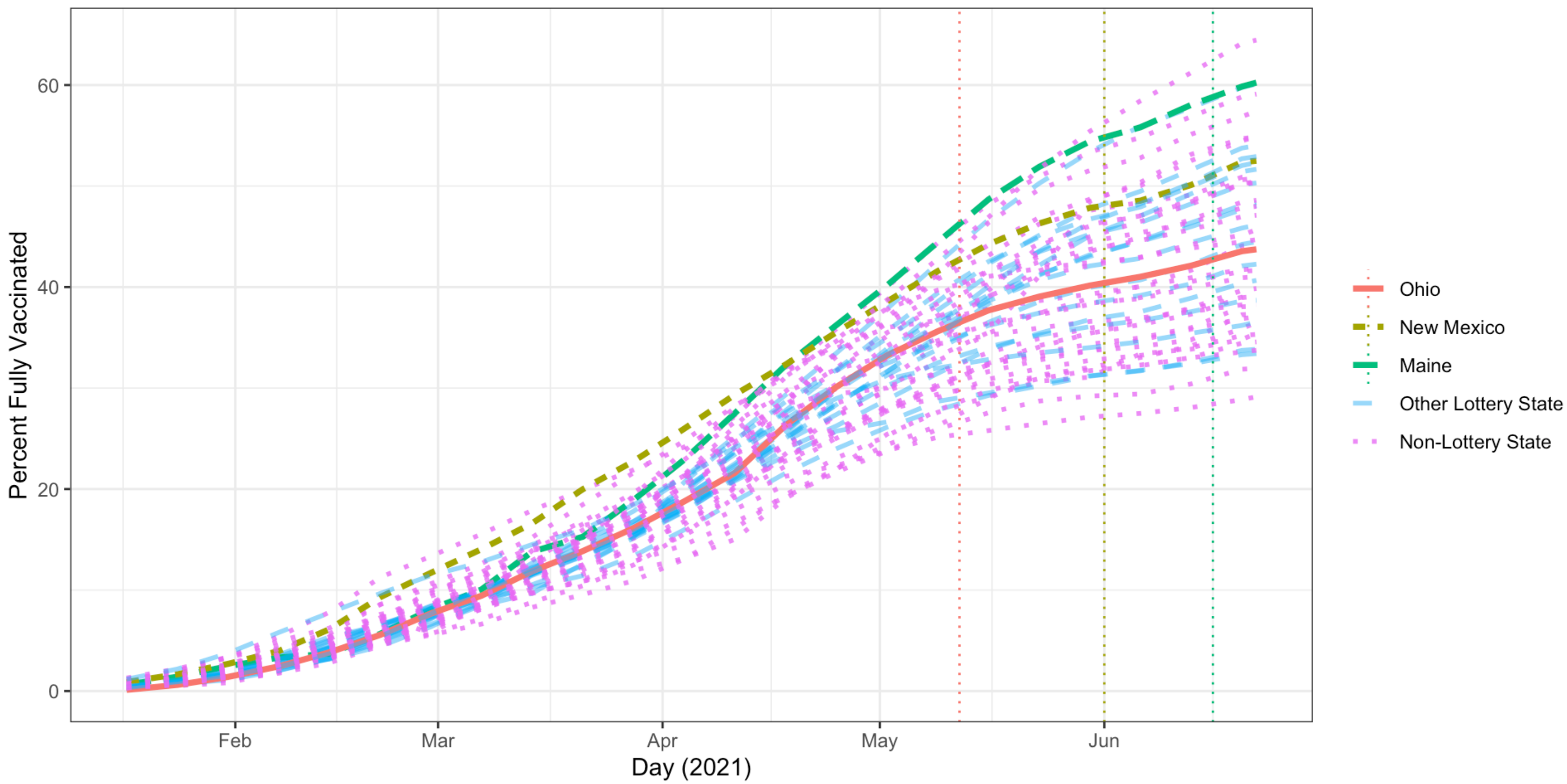
- Allow for staggered adoption
- Reduce interpolation
- Account for exposed series outside of the convex hull

Key Ideas

- Allow for staggered adoption
- Reduce interpolation
- Account for exposed series outside of the convex hull
- Incorporate control series on different outcomes/scales from treated series

Motivating Example

We saw in the previous example that multiple states initiated lotteries at different times.



Plot of percent fully vaccinated rates by U.S. state, Jan.-Sept. 2021

Methods

Standard SCM

The standard SCM had weights w_i that minimized:

$$\sum_{k=1}^{K} v_k \left(X_{1k} - \left(\sum_{i=1}^{n} X_{0ik} w_i \right) \right)^2,$$

for covariates $k = 1, \ldots, K$.

General: Penalized SCM

This can be extended with a penalty term ξ on a function of the weights $f(w_i)$ to minimize:

$$\sum_{k=1}^{K} v_k \left(X_{1k} - \left(\sum_{i=1}^{n} X_{0ik} w_i \right) \right)^2 + \xi \sum_{i=1}^{n} f(w_i).$$

The penalty can reduce interpolation (force closer matches to specific units) or reduce discrepancy from an outcome model.

De-Meaned SCM

BALANCING, REGRESSION, DIFFERENCE-IN-DIFFERENCES AND SYNTHETIC CONTROL METHODS:
A SYNTHESIS

Nikolay Doudchenko
Guido W. Imbens

Synthetic controls with imperfect pretreatment fit

BRUNO FERMAN
Sao Paulo School of Economics-FGV

CRISTINE PINTO
Sao Paulo School of Economics-FGV

De-Meaned SCM

Idea

De-mean the pre-treatment data, fit SC to the de-meaned observations, and apply the weights to both the post-treatment time trends and levels.

Incorporates idea of diff-in-diff of focusing on matching trends instead of levels.

De-Meaned SCM

Idea

De-mean the pre-treatment data, fit SC to the de-meaned observations, and apply the weights to both the post-treatment time trends and levels.

Incorporates idea of diff-in-diff of focusing on matching trends instead of levels.

Warning

Matching pre-treatment trends may not lead to stable weights going forward.

Different interpretation of weights.

Synthetic DID

Synthetic Difference-in-Differences†

By Dmitry Arkhangelsky, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager*

SDID: Procedure

Idea

Incorporate unit weighting of SC with unit fixed effects of DID and time weighting. “Localized” TWFE model.

SDID: Procedure

Idea

Incorporate unit weighting of SC with unit fixed effects of DID and time weighting. "Localized" TWFE model.

1. Compute regularization parameter

SDID: Procedure

> **Idea**
>
> Incorporate unit weighting of SC with unit fixed effects of DID and time weighting. "Localized" TWFE model.

1. Compute regularization parameter
2. Compute regularized, intercept-adjusted/de-meaned SC weights

SDID: Procedure

> **Idea**
>
> Incorporate unit weighting of SC with unit fixed effects of DID and time weighting. "Localized" TWFE model.

1. Compute regularization parameter
2. Compute regularized, intercept-adjusted/de-meaned SC weights
3. Compute time weights

SDID: Procedure

Idea

Incorporate unit weighting of SC with unit fixed effects of DID and time weighting. "Localized" TWFE model.

1. Compute regularization parameter
2. Compute regularized, intercept-adjusted/de-meaned SC weights
3. Compute time weights
4. Conduct weighted TWFE regression

Augmented Synthetic Control

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2021, VOL. 116, NO. 536, 1789–1803: Theory and Methods Special Section on Synthetic Control Methods
https://doi.org/10.1080/01621459.2021.1929245

The Augmented Synthetic Control Method

Eli Ben-Michael[a], Avi Feller[b], and Jesse Rothstein[c]

Augmented Synthetic Control

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2021, VOL. 116, NO. 536, 1789–1803: Theory and Methods Special Section on Synthetic Control Methods
https://doi.org/10.1080/01621459.2021.1929245

The Augmented Synthetic Control Method

Eli Ben-Michael[a], Avi Feller[b], and Jesse Rothstein[c]

> **Idea**
>
> De-bias SC estimate using an outcome model for the time series.

Ridge-Adjusted ASCM: Procedure

Ridge-Adjusted ASCM: Procedure

1. Compute SC weights

Ridge-Adjusted ASCM: Procedure

1. Compute SC weights
2. Fit outcome model (i.e., ridge regression) to control data (post-treatment outcomes ~ pre-treatment outcomes)

Ridge-Adjusted ASCM: Procedure

1. Compute SC weights
2. Fit outcome model (i.e., ridge regression) to control data (post-treatment outcomes ~ pre-treatment outcomes)
3. Get model estimates for all units' post-treatment outcomes

Ridge-Adjusted ASCM: Procedure

1. Compute SC weights
2. Fit outcome model (i.e., ridge regression) to control data (post-treatment outcomes ~ pre-treatment outcomes)
3. Get model estimates for all units' post-treatment outcomes
4. Find discrepancy between model estimate for treated unit and model estimate for synthetic unit (weighted avg of model estimates for control units)

Ridge-Adjusted ASCM: Procedure

1. Compute SC weights
2. Fit outcome model (i.e., ridge regression) to control data (post-treatment outcomes ~ pre-treatment outcomes)
3. Get model estimates for all units' post-treatment outcomes
4. Find discrepancy between model estimate for treated unit and model estimate for synthetic unit (weighted avg of model estimates for control units)
5. Add this difference to SC estimator

ASCM: Additional Details

ASCM: Additional Details

- If pre-treatment fit is good, discrepancy will be small and adjustment will have little effect

ASCM: Additional Details

- If pre-treatment fit is good, discrepancy will be small and adjustment will have little effect
- Allows for level shift by capturing a consistent discrepancy

ASCM: Additional Details

- If pre-treatment fit is good, discrepancy will be small and adjustment will have little effect
- Allows for level shift by capturing a consistent discrepancy
- Can still express as weights of control units, but negative weights now allowed; penalizes discrepancy from SC weights

ASCM: Multiple Periods

Synthetic controls with staggered adoption

Eli Ben-Michael[1] | Avi Feller[2] | Jesse Rothstein[2]

Options:

- Fit separate SCM for each unit
- Fit SCM on average of treated units
- Partially pooled SCM: mix of both
- With intercepts, similar trade-offs to weighted DID approaches

SDID and ASCM: Summary

Statistical Methods in Medical Research
1–16

DOI: 10.1177/09622802231224638
journals.sagepub.com/home/smm
Sage

The augmented synthetic control method in public health and biomedical research

Taylor Krajewski[1] and Michael Hudgens[1]

SDID and ASCM: Summary

The augmented synthetic control method in public health and biomedical research

Statistical Methods in Medical Research
1–16

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09622802231224638
journals.sagepub.com/home/smm
Sage

Taylor Krajewski[1] and Michael Hudgens[1]

Advantages:

- Allow intercept shift
- Allow negative weights with some sparsity
- Improved performance in settings with poor SC fit

SDID and ASCM: Summary

The augmented synthetic control method in public health and biomedical research

Statistical Methods in Medical Research
1–16

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09622802231224638
journals.sagepub.com/home/smm
Sage

Taylor Krajewski[1] and Michael Hudgens[1]

Disadvantages:

- Allow extrapolation
- Loss of interpretation of weights
- More user degrees of freedom
- Challenging inference

Generalized Synthetic Control

Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models

Yiqing Xu

Idea: “Interactive Fixed Effects”

Use control unit data to estimate unit fixed effects and some set of unknown time-varying coefficients (factors).

Use these coefficients to estimate treated unit fixed effects.

Use the treated unit FEs and time-varying effects to estimate counterfactuals for treated unit-periods.

Matrix Completion

- Similar to interactive fixed effects
- Uses a continuous penalty instead of a sparsity-inducing penalty on the unknown time-varying coefficients.

GSC, IFE, and Matrix Completion

Advantages:

- Can achieve better fit
- Accomodates staggered adoption and general time-varying treatment
- Allows quick and efficient estimation for multiple treated units

Disadvantages/Assumptions:

- Requires a fixed (but unknown) set of time-varying factors across time and units
- Number of factors usually selected by cross-validation
- Allows extrapolation and loses strict weighting interpretability

Bayesian Structural Time Series Modeling

INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS

BY KAY H. BRODERSEN, FABIAN GALLUSSER, JIM KOEHLER, NICOLAS REMY AND STEVEN L. SCOTT

Idea

Combine three information sources in state-time model:

Bayesian Structural Time Series Modeling

INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS

BY KAY H. BRODERSEN, FABIAN GALLUSSER, JIM KOEHLER, NICOLAS REMY AND STEVEN L. SCOTT

Idea

Combine three information sources in state-time model:

- Bayesian priors on covariate importance

Bayesian Structural Time Series Modeling

INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS

BY KAY H. BRODERSEN, FABIAN GALLUSSER, JIM KOEHLER, NICOLAS REMY AND STEVEN L. SCOTT

Idea

Combine three information sources in state-time model:

- Bayesian priors on covariate importance
- Time series modeling of outcome pre-intervention

Bayesian Structural Time Series Modeling

INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS

BY KAY H. BRODERSEN, FABIAN GALLUSSER, JIM KOEHLER, NICOLAS REMY AND STEVEN L. SCOTT

Idea

Combine three information sources in state-time model:

- Bayesian priors on covariate importance
- Time series modeling of outcome pre-intervention
- State-space model for treated unit based on controls

BSTS: Summary

- Allows extrapolation: weights do not sum to 1
- Incorporates prior information
- Explicit time series modeling

Estimating the population-level impact of vaccines using synthetic controls

Christian A. W. Bruhn[a], Stephen Hetterich[b], Cynthia Schuck-Paim[b], Esra Kürüm[a,c], Robert J. Taylor[b], Roger Lustig[b], Eugene D. Shapiro[a,d], Joshua L. Warren[a,e], Lone Simonsen[b,f,g], and Daniel M. Weinberger[a,1]

Evaluating the Impact of Meningococcal Vaccines With Synthetic Controls

Ottavia Prunas, Daniel M. Weinberger, Duccio Medini*, Michele Tizzoni, and Lorenzo Argante

Tradeoffs and Interpretations

Validity: Interpolation vs. Extrapolation

Validity: Interpolation vs. Extrapolation

- Pooling across units enables better fit, but loses unit-specific information

Validity: Interpolation vs. Extrapolation

- Pooling across units enables better fit, but loses unit-specific information
- Extrapolation allows better pre-treatment fit, but may over-fit or rely on additional assumptions

Validity: Interpolation vs. Extrapolation

- Pooling across units enables better fit, but loses unit-specific information
- Extrapolation allows better pre-treatment fit, but may over-fit or rely on additional assumptions
- Reducing interpolation is crucial for some settings (non-linear outcomes)

Generalizability-Bias-Variance Tradeoffs

Generalizability-Bias-Variance Tradeoffs

- Advanced methods allow use of more control series: reduces variance but may introduce bias

Generalizability-Bias-Variance Tradeoffs

- Advanced methods allow use of more control series: reduces variance but may introduce bias
- Cross-validation procedures can improve fit but reduce efficiency

Generalizability-Bias-Variance Tradeoffs

- Advanced methods allow use of more control series: reduces variance but may introduce bias
- Cross-validation procedures can improve fit but reduce efficiency
- Accounting for heterogeneities improves generalizability and reduces bias but increases variance

Generalizability-Bias-Variance Tradeoffs

- Advanced methods allow use of more control series: reduces variance but may introduce bias
- Cross-validation procedures can improve fit but reduce efficiency
- Accounting for heterogeneities improves generalizability and reduces bias but increases variance
- Still require common factors, exogenous shocks, no intervening treatments

Interpretability of Weights

Interpretability of Weights

- A key benefit of SC is its interpretability

Interpretability of Weights

- A key benefit of SC is its interpretability
- This is somewhat lost in more advanced approaches

Interpretability of Weights

- A key benefit of SC is its interpretability
- This is somewhat lost in more advanced approaches
- The interpretability is tied to justifying the assumptions as well

Questions?