Synthetic Control Under Interference: Detecting and Correcting Bias

Joao Alipio-Correa

Political Science & Statistics | University of Pittsburgh

SCM in Political Science

SCM emerged as an important tool for analyzing rare political events:

- **Civil wars:** Coercion, governance, and political behavior in civil war. *Journal of Peace Research*, 2024
- **Polarization:** Partisan Enclaves and Information Bazaars: Mapping Selective Exposure to News. *Journal of Politics*, 2022
- **Far Right:** Do Voters Polarize When Radical Parties Enter Parliament? *American Journal of Political Science*, 2019
- **Religion & Politics:** Government Religious Discrimination, Support of Religion, and Societal Violence in Western Democracies. *Comparative Political Studies*, 2024
- **Political Economy:** From Rents to Welfare: Why Are Some Oil-Rich States Generous to Their People? *American Political Science Review*, 2024
- **Regimes:** The Rush to Personalize: Power Concentration after Failed Coups in Dictatorships. *British Journal of Political Science*, 2023
- **Institutional change:** Comparative politics and the synthetic control method. *American Journal of Political Science*, 2015

Causal Inference and Interference

When policies, conflicts, or shocks *spill over* to neighboring regions, do we still have valid donor pools under Synthetic Control?

Outline

What is the Synthetic Control Method (SCM)?

- Enables inference with a small number (or single) treated units;
- Build a synthetic version of the treated unit as a counterfactual weighting unaffected units.
- Potential outcomes for treated unit:
  - Y_{1t}^N: Outcome in absence of intervention (counterfactual).
  - Y_{1t}^I: Outcome under intervention.
- Treatment effect:

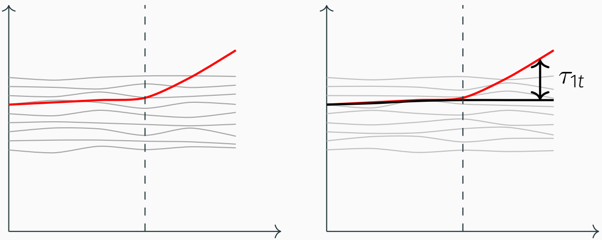
$$\tau_{1t} = Y_{1t}^I - Y_{1t}^N, \quad t > T_0.$$

SCM: How It Works

$$\hat{Y}_{1t}^{N} = \sum_{j=2}^{J+1} w_j Y_{jt}, \quad t > T_0.$$

- Optimal weights W^*: Minimize discrepancy in pre-treatment characteristics and $\|\cdot\|_V$ reflects predictors importance:

$$W^* = \arg\min_{W} \|X_1 - X_0 W\|_V,$$



SCM and SUTVA

- **Stable Unit Treatment Value Assumption (SUTVA):**

$$Y_{it}(Z_i, Z_{-i}) \;=\; Y_{it}(Z_i) \quad \forall i$$

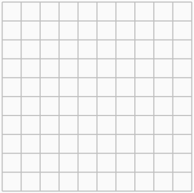
No interference: No unit's outcome depends on other units' treatment status.

- **Crucial Assumption:** The donor units remain *untreated*. Any violation (e.g., partial exposure) can bias the synthetic estimate.
- **SUTVA violation:** Suppose donor j receives an interference term δ_{jt}. The synthetic counterfactual becomes

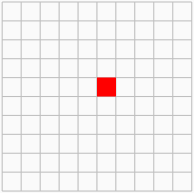
$$\hat{Y}_{it}^{N} \;=\; \sum_{j \neq i} w_j \,(Y_{jt}^{N} + \delta_{jt}),$$

so the estimated effect $\hat{\tau}_{it}$ deviates by $\sum_j w_j \delta_{jt}$ from the *true* τ_{it}.

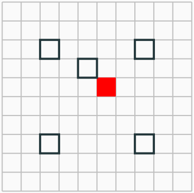
Stages of SCM Construction



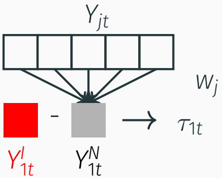
1: Units



2: Single Treated Unit

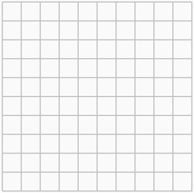


3: Units for Synthetic Control

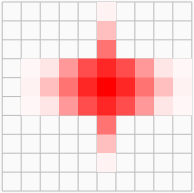


4: Treatment Effect

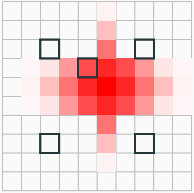
Stages of SCM Construction



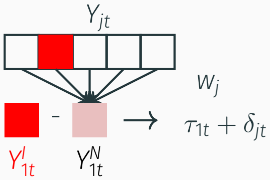
1: Units



2: Treatment diffusion

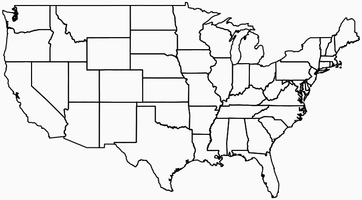


3: Units for Synthetic Control



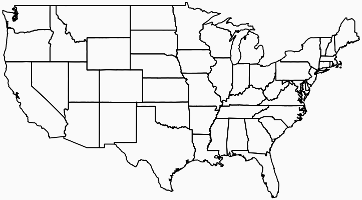
4: Contaminated Treatment Effect

Simulated data

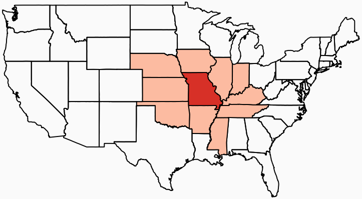


Units map

Simulated data

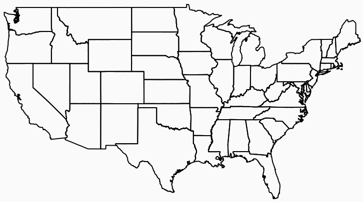


Units map

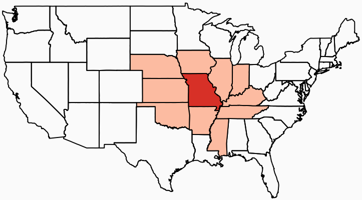


Missouri being treated

Simulated data



Units map

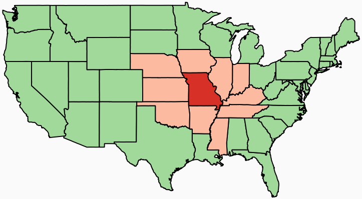


Missouri being treated

Simulated data for an intervention in Missouri with true ATT $\tau = 4$ and interfering the outcome for nearby units by a parameter of $\rho = 0.6$

Closer units are more affected by interference than farther away ones. But how can we compare and test if this interference is at play?

Contrast setup



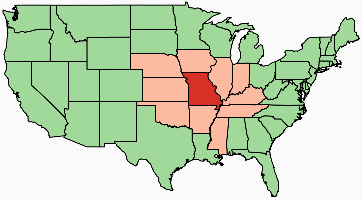
Contrast for Missouri

Let $i \in \mathcal{U} = \{1, \ldots, N\}$ index units (in this case, US states)

Fix the treated unit ($p \in \mathcal{U}$) at the center and compute distances d_{ip} partitioning the space in non-overlapping rings

$$c_0 < c_1 < \cdots < c_K$$

Contrast setup



Contrast for Missouri

Let $i \in \mathcal{U} = \{1, \ldots, N\}$ index units (in this case, US states)

Fix the treated unit ($p \in \mathcal{U}$) at the center and compute distances d_{ip} partitioning the space in non-overlapping rings

$$c_0 < c_1 < \cdots < c_K$$

Each ring being identified as:

$$r_{ip} = k \iff c_{k-1} \leq d_{ip} < c_k, \quad k = 1, \ldots, K$$

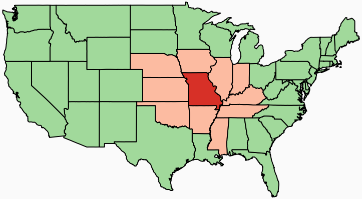
Then assign units to fully disjoint rings according to their distance from p:

- Focus ring: $R_A \subset \{1, \ldots, Q\}$
- Comparison ring: $R_B \subset \{Q+1, \ldots, K\}$

And define groups:

- $A_p = \{i \neq p : r_{ip} \in R_A\}$
- $B_p = \{i \neq p : r_{ip} \in R_B\}$

Contrast setup - Z value



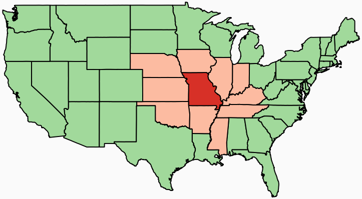
But what are we comparing?

Let $t \in \mathcal{T}$ index time, T_0 be the treatment period for unit p, and Y_{it} represent the outcome

Define two disjoint sets of periods for each window w:

$$\mathcal{T}_w^{\text{pre}}, \mathcal{T}_w^{\text{post}} \subset \mathcal{T}, \quad \mathcal{T}_w^{\text{pre}} \cap \mathcal{T}_w^{\text{post}} = \emptyset$$

Contrast setup - Z value



But what are we comparing?

Let $t \in \mathcal{T}$ index time, T_0 be the treatment period for unit p, and Y_{it} represent the outcome

Define two disjoint sets of periods for each window w:

$$\mathcal{T}_w^{\text{pre}}, \mathcal{T}_w^{\text{post}} \subset \mathcal{T}, \quad \mathcal{T}_w^{\text{pre}} \cap \mathcal{T}_w^{\text{post}} = \emptyset$$

And set windows of interest for the difference in outcome, such as:

w	$\mathcal{T}_w^{\text{pre}}$	$\mathcal{T}_w^{\text{post}}$
full	$\{t < T_0\}$	$\{t > T_0\}$
year-1	$\{T_0 - 1\}$	$\{T_0 + 1\}$
sym-n	$\{T_0 - n, \dots, T_0 - 1\}$	$\{T_0 + 1, \dots, T_0 + n\}$

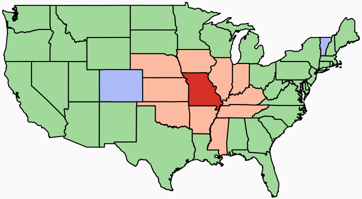
And for every unit i and window w, define a difference-in-means statistic:

$$Z_i^{(w)} = \bar{Y}_{i,\text{post}(w)} - \bar{Y}_{i,\text{pre}(w)}$$

where: $$\bar{Y}_{i,\text{post}(w)} = \frac{1}{|\mathcal{T}_w^{\text{post}}|} \sum_{t \in \mathcal{T}_w^{\text{post}}} Y_{it}$$

and $$\bar{Y}_{i,\text{pre}(w)} = \frac{1}{|\mathcal{T}_w^{\text{pre}}|} \sum_{t \in \mathcal{T}_w^{\text{pre}}} Y_{it}$$

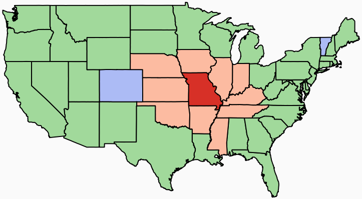
Contrast setup - first test



$Z_i^{(w)} \to$ average outcome variation for each i between post-pre periods in window w.

Anomalous values in units nearby the treated hint at potential interference

Contrast setup - first test



$Z_i^{(w)} \rightarrow$ average outcome variation for each i between post-pre periods in window w.

Anomalous values in units nearby the treated hint at potential interference

state	$Z^{(\text{full})}$	$Z^{(\text{year-1})}$	$Z^{(\text{sym-3})}$
Missouri	4.0066	3.9159	3.9381
Iowa	2.3640	2.4193	2.3539
Colorado	-0.0414	-0.1069	0.0060
Vermont	0.02501	-0.1115	-0.0886

For each window w, collect $Z_i^{(w)}$ for $i \in A_p$ and $Z_i^{(w)}$ for $i \in B_p$, and let

$$\bar{Z}_{A_p}^{(w)} = \frac{1}{|A_p|} \sum_{i \in A_p} Z_i^{(w)}, \quad \bar{Z}_{B_p}^{(w)} = \frac{1}{|B_p|} \sum_{i \in B_p} Z_i^{(w)}$$

denote the group means for each ring set and build:

$$t_p = \frac{\bar{Z}_{A_p} - \bar{Z}_{B_p}}{\sqrt{s_P^2 \left(\frac{1}{|A_p|} + \frac{1}{|B_p|} \right)}}$$

Large $|t_p| \Rightarrow$ evidence that proximity ring(s) differ in mean outcome change relative to farther rings

Contrast setup - randomization

Checking whether average for nearby units	$\neq$	units farther away from treated unit (around treatment)	✓

Can we reject the null of no interference?

Contrast setup - randomization

Checking whether average for nearby units	$\neq$	units farther away from treated unit (around treatment)	✓

Can we reject the null of no interference?

Randomization inference:

$$H_0 : \left\{ Z_i^{(w)} \right\}_{i \in \mathcal{U}} \text{ is invariant to which unit is labelled "treated".}$$

i.e.: Pattern of interference around treated unit is no different than the pattern around any other unit in the space

Contrast setup - randomization II

Algorithm

1. Compute t_p for every $p \in \mathcal{U}$ as above.
2. Let t_0 be the statistic for the **actual treated unit** $p = p^\star$.
3. Exact two-sided p-value:

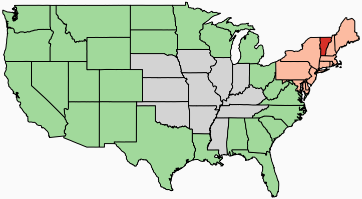
$$\hat{p} = \frac{1 + \sum_{p \in \mathcal{U}} \mathbf{1}(|t_p| \geq |t_0|)}{N + 1}$$

Contrast setup - randomization II

Algorithm

1. Compute t_p for every $p \in \mathcal{U}$ as above.
2. Let t_0 be the statistic for the **actual treated unit** $p = p^\star$.
3. Exact two-sided p-value:

$$\hat{p} = \frac{1 + \sum_{p \in \mathcal{U}} \mathbf{1}(|t_p| \geq |t_0|)}{N + 1}$$



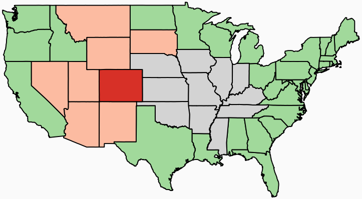
Contrast for Vermont

Contrast setup - randomization II

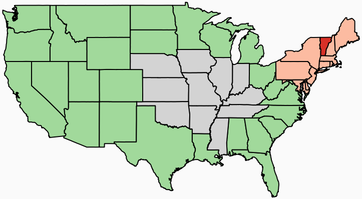
Algorithm

1. Compute t_p for every $p \in \mathcal{U}$ as above.
2. Let t_0 be the statistic for the **actual treated unit** $p = p^\star$.
3. Exact two-sided p-value:

$$\hat{p} = \frac{1 + \sum_{p \in \mathcal{U}} \mathbf{1}(|t_p| \geq |t_0|)}{N + 1}$$



Contrast for Colorado



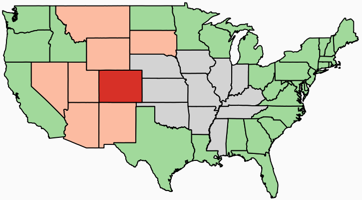
Contrast for Vermont

Contrast setup - randomization II

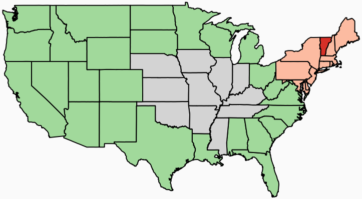
Algorithm

1. Compute t_p for every $p \in \mathcal{U}$ as above.
2. Let t_0 be the statistic for the **actual treated unit** $p = p^\star$.
3. Exact two-sided p-value:

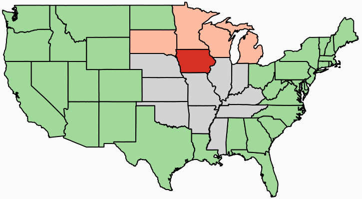
$$\hat{p} = \frac{1 + \sum_{p \in \mathcal{U}} \mathbf{1}(|t_p| \geq |t_0|)}{N + 1}$$



Contrast for Colorado



Contrast for Vermont



Contrast for Iowa

Contrast setup- randomization II

Algorithm

1. Compute t_p for every $p \in \mathcal{U}$ as above.
2. Let t_0 be the statistic for the **actual treated unit** $p = p^\star$.
3. Exact two-sided p-value:

$$\hat{p} = \frac{1 + \sum_{p \in \mathcal{U}} \mathbf{1}(|t_p| \geq |t_0|)}{N + 1}$$

state	t_p	A_p	B_p
MO	4.4207	AR, IL, IN, ...	AL, AZ, CA, ...
VT	-0.2169	CT, DE, ME, ...	AL, AZ, CO, ...
CO	0.3428	AZ, MT, NV, ...	AL, CA, CT, ...
IA	-0.3312	MI, MN, SD, ...	AL, AZ, CA, ...

And from this simulated scenario we obtained p-value = 0.0408

Contrast setup - alternative contrasts

Where does it end?

Detecting whether interference is present ✓

Detecting where interference is no longer statistically significant:

Contrast setup - alternative contrasts

Where does it end?

Detecting whether interference is present ✓

Detecting where interference is no longer statistically significant:

Instead of contrasting

$A_{p^\star} = \{i \neq p^\star : r_{ip^\star} = 1\}$ vs.

$B_{p^\star} = \{i \neq p^\star : r_{ip^\star} \in \{2, 3, 4, 5\}$

to obtain the standard $t_{p^\star}^{(\text{1 vs 2:5})}$

Contrast: $A_{p^\star} = \{i \neq p^\star : r_{ip^\star} = 2\}$ vs.

$B_{p^\star} = \{i \neq p^\star : r_{ip^\star} \in 3\} \rightarrow t_{p^\star}^{(\text{2 vs 3})}$

Contrast setup - alternative contrasts

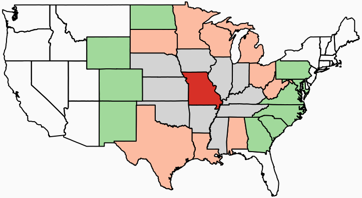
Where does it end?

Detecting whether interference is present ✓

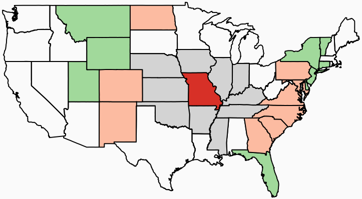
Detecting where interference is no longer statistically significant:

Instead of contrasting
$A_{p^\star} = \{i \neq p^\star : r_{ip^\star} = 1\}$ vs.
$B_{p^\star} = \{i \neq p^\star : r_{ip^\star} \in \{2,3,4,5\}$
to obtain the standard $t_{p^\star}^{(1 \text{ vs } 2:5)}$

Contrast: $A_{p^\star} = \{i \neq p^\star : r_{ip^\star} = 2\}$ vs.
$B_{p^\star} = \{i \neq p^\star : r_{ip^\star} \in 3\} \to t_{p^\star}^{(2 \text{ vs } 3)}$



2 vs 3 Contrast for Missouri, p = 0.9591



3 vs 4 Contrast for Colorado, p = 0.5102041

Interference Confirmed. Now What?

Interference ✓

Two options:

- *1. Keeping them unmodified* leads to biased synthetic estimates.
- *2. Simply dropping* suspect donors might degrade the pre-treatment match.

Interference Confirmed. Now What?

Interference ✓

Two options:

- *1. Keeping them unmodified* leads to biased synthetic estimates.
- *2. Simply dropping* suspect donors might degrade the pre-treatment match.

2.1 But at least now we are able to make an informed decision on which units to drop

Interference Confirmed. Now What?

Interference ✓

Two options:

- *1. Keeping them unmodified* leads to biased synthetic estimates.
- *2. Simply dropping* suspect donors might degrade the pre-treatment match.

    2.1 But at least now we are able to make an informed decision on which units to drop

- *3. Adjust for it:* Use a secondary set of weights to attenuate contamination in the donor pool

    Spatial reach measure as the weights

Spatial Reach: A Continuous Proximity Index

- For donor j, let d_j be its distance to the treated unit.

$$\mathrm{SR}_j = \frac{1}{1 + \exp\left[-\kappa(d_j - c)\right]},$$

  - c is typically the *mean* or *median* distance to center the logistic curve.
  - κ scales how steeply SR_j transitions from near 0 to near 1.

- **Parameter Tuning:** κ trimmed between the 2.5% and 97.5% percentiles of $\{d_j\}$, ensuring a smooth but complete range.
- **Interpretation:** $\mathrm{SR}_j \approx 0$ if donor j is very close, and ≈ 1 if it is far.

Bias Correction Strategies

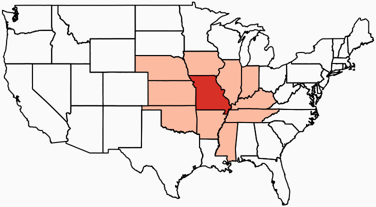
Solution	Optimization	Simplex	Consequence
Rescaling	$\min_w \lVert \mathbf{X}_1 - \mathbf{X}_0^* \, w_j \rVert^2$ with $X^*_{k,j} = X_{k,j} \times \mathrm{SR}_j$	✓	Downweights exposed units; Retains convex weights
Ridge constrained	$\min_w \lVert \mathbf{X}_1 - \mathbf{X}_0 \, w_j \rVert^2 + \lambda \sum_j \mathrm{SR}_j \, w_j^2$	✓	Penalize large SCM weights Moderate contamination
Ridge unconstrained	$\min_w \lVert \mathbf{X}_1 - \mathbf{X}_0 \, w_j \rVert^2 + \lambda \sum_j \mathrm{SR}_j \, w_j^2$	✗	Allows negative SCM weights Aggressively offset contamination

Simplex constraint: $w_j \geq 0, \ \sum_j w_j = 1$

- Units are only allowed to have positive weights
- Unit weights add up to 1

US Simulation

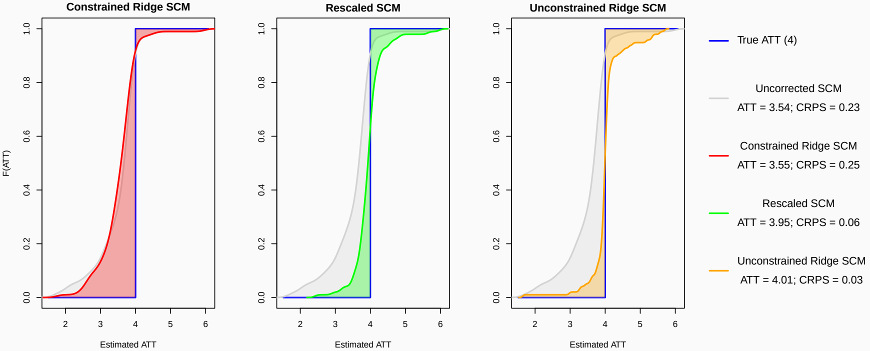
Setup: Intervention in Missouri with true effect size $\tau = 4$ and spillover intensity $\rho = 0.6$.



Compare the uncorrected biased SCM versus the three correction approaches

Metrics: Bias in the estimated ATT, pre-treatment RMSE, and CRPS.

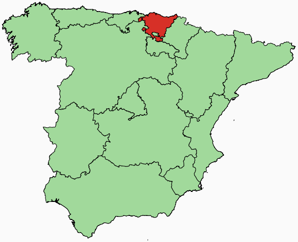
US Simulation results



Simulation under $\tau = 4$ and $\rho = 0.6$

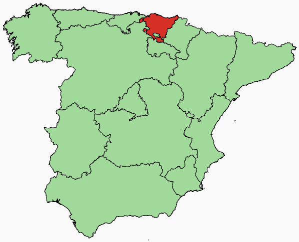
Consistent across all effect sizes τ and spillover intensity ρ

Interference in Applied Research

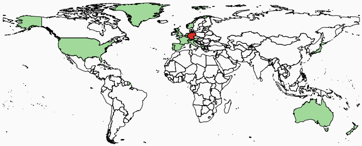


Abadie et al (2003) Conflict in the Basque
p = 0.22

Interference in Applied Research

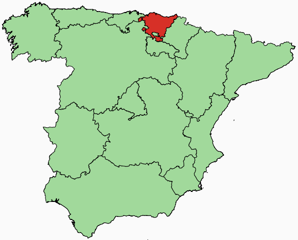


Abadie et al (2003) Conflict in the Basque
p = 0.22

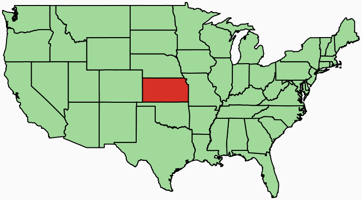


Abadie et al (2015) German Reunification
p = 0.46

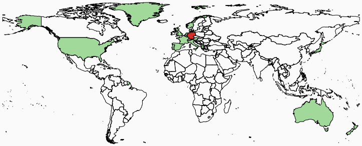
Interference in Applied Research



Abadie et al (2003) Conflict in the Basque
p = 0.22

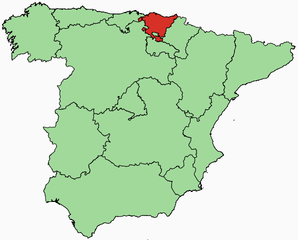


Ben-Michael et al (2021) Kansas tax cut
p = 0.18

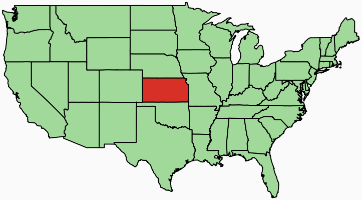


Abadie et al (2015) German Reunification
p = 0.46

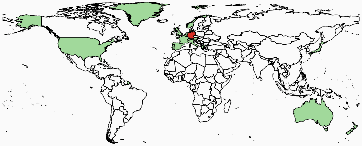
Interference in Applied Research



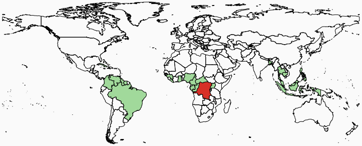
Abadie et al (2003) Conflict in the Basque
p = 0.22



Ben-Michael et al (2021) Kansas tax cut
p = 0.18



Abadie et al (2015) German Reunification
p = 0.46



Kikuta (2020); Civil war and deforestation
p = 0.33

Interference in Applied Research

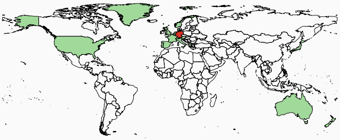
Application	Coverage	Interference
Abadie et al (2003)	✓	✗
Ben-Michael et al (2021)	✓	✗
Abadie et al (2015)	✗	✗
Kikuta (2019)	✗	✗

Interference in Applied Research

Application	Coverage	Interference
Abadie et al (2003)	✓	✗
Ben-Michael et al (2021)	✓	✗
Abadie et al (2015)	✗	✗
Kikuta (2019)	✗	✗
Expanded German Reunification	✓	✓

Interference in Applied Research

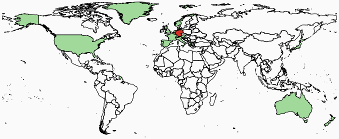
Application	Coverage	Interference
Abadie et al (2003)	✓	✗
Ben-Michael et al (2021)	✓	✗
Abadie et al (2015)	✗	✗
Kikuta (2019)	✗	✗
Expanded German Reunification	✓	✓



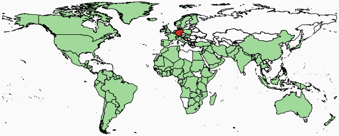
Abadie et al (2015) German Reunification
p = 0.46

Interference in Applied Research

Application	Coverage	Interference
Abadie et al (2003)	✓	✗
Ben-Michael et al (2021)	✓	✗
Abadie et al (2015)	✗	✗
Kikuta (2019)	✗	✗
Expanded German Reunification	✓	✓



Abadie et al (2015) German Reunification
$p = 0.46$



Expanded German Reunification
$p = 0.016$

Interference in Applied Research

Researchers try to address SUTVA violations and patterns of interference by removing units $\rightarrow$ results conditioned on contagion

Risk $\rightarrow$ dropping too many units

Under Potential Outcomes, the DGP and a suitable identification strategy depends on: empirics *AND* how the missing potential outcome is set up

- In the SCM case: which units are in the donor pool

Replication Examples

Comparative politics and the synthetic control method (Abadie, Diamond, & Hainmueller, 2015): German Reunification

Approach	Metric	Germany
Base	ATT Pre-RMSE	-1549.9 119.08
Rescaled	ATT Pre-RMSE	-1601.5 279.03
Penalized, Constrained *	ATT Pre-RMSE	-1103.4 80.43
Penalized, Unconstrained *	ATT Pre-RMSE	136.1 59.5

Rescaling adjusted for contamination → larger effect

Constrained Ridge adjust for contamination and large weights → attenuation

Unconstrained Ridge extrapolate simplex for aggressive correction → reversal

Concluding remarks - Detection

A) Detection

- **Coverage**: Ensure proper donor units coverage to compose the missing potential outcome;
- **Detection test**: Using randomization inference, assess whether interference is at place in the empirical setting;
- **Alternative contrast**: By adapting the contrast, identify where interference is no longer detected;
- **Detect Interference First**: If no violation is detected, standard SCM suffices;

Concluding remarks - Correction

B) Correction

- **SR weight**: If interference $\rightarrow$ subject the SCM optimization problem to network-specific weights;
- **Minor to moderate interference**: Rescaling or Constrained Ridge can mitigate moderate bias while retaining the notion of a convex combination.;
- **Severe Interference**: Unconstrained Ridge achieves lower bias at the cost of extrapolating out of the simplex;

Ongoing Extensions

- **Inverse Propensity Weighting for Rescaling Approach**

  HT–Hájek Spatial Weights

  Spatial–reach $f(d)$ as *propensity to avoid spillover*: $\pi_i = 1 - f(d_{iD})$

  Use stabilized Horvitz–Thompson weights $w_i = \dfrac{1/\pi_i}{\sum_j 1/\pi_j}$ inside SCM

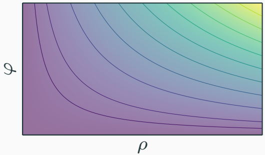
- **Multiple Comparison & Dynamic Networks**

- **Sensitivity to Interference**

  Inject controlled spillovers in outcomes & covariates: intensity $\rho \in [0, 1]$, decay φ

  Re-run SCM over a (ρ, φ) grid; track standardized shift

  Contours show ATT shift required to overturn conclusions



(Lighter $\rightarrow$ larger ATT shift)