Panel Methods (3)

Synthetic Control and Extensions

Yiqing Xu (Stanford University)

July 31st, 2024

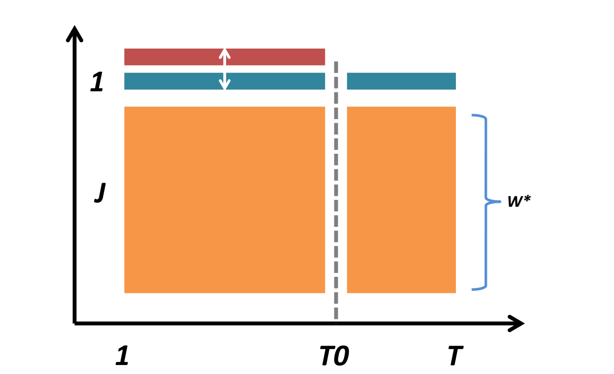
This Lecture

- What can we do when the parallel trends assumption is violated?
- Solution 1: conditional on pre-treatment covariates $\rightarrow$ semi-parametric DID
- Solution 2: conditional on pre-treatment covariates & outcomes $\rightarrow$ synthetic control method
- This lecture
  - The synthetic control method
  - The latent factor approach
  - Doubly-robust methods

The Synthetic Control Method (SCM)

SCM: Basic Idea

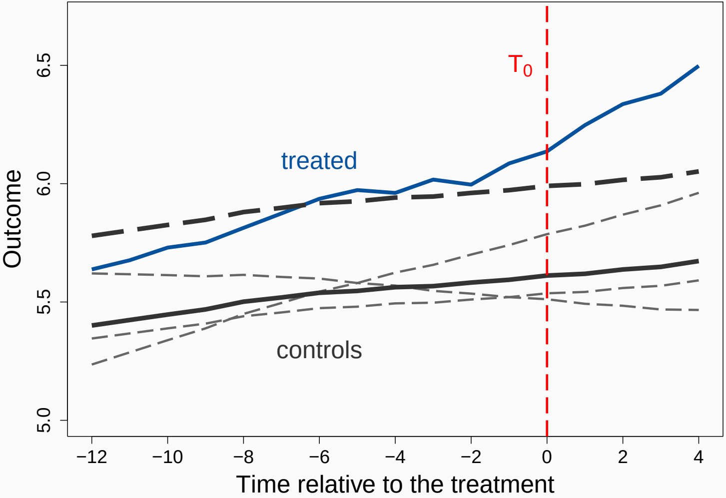
- $J + 1$ units in periods $1, 2, \ldots, T$; one treated "1", J controls
- Region "1" is exposed to the intervention after period T_0
- We aim to estimate the effect of the intervention on Region "1"



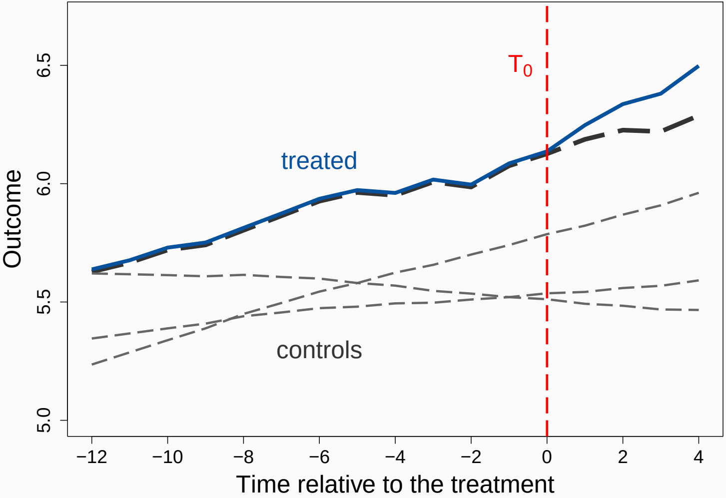
SCM: Insights

- Athey and Imbens (2016): "[a]rguably the most important innovation in the policy evaluation literature in the last 15 years."
- A combo of many innovations
    - Take advantage of pre-treatment outcomes
    - Use cross-sectional instead of temporal correlations in data
    - Construct a convex combination of donors to construct a counterfactual
    - Reserve some pre-treatment periods for testing

Difference-in-Differences (DID)



SCM (and Many Extensions)



Theoretical Justification

$$Y_{it} = \tau_{it} D_{it} + \theta_t' Z_i + \xi_t + \lambda_i' f_t + \varepsilon_{it}$$

or

$$\begin{cases} Y_{it}(0) & = \theta_t' Z_i + \xi_t + \lambda_i' f_t + \varepsilon_{it} \\ Y_{it}(1) & = Y_{it}^0 + \tau_{it} \end{cases}$$

- Suppose there are R time-varying signals f_t out there
- Each unit (e.g. country, participant) picks up a fixed linear combination of these signals based on factor loadings λ_i
- Since these "confounders" are evidenced in the pre-treatment outcomes for both treated and controls, we can try to use this information to "balance on" these confounders
- We will discuss the model-based approach later

Theoretical Justification

$$\begin{cases} Y_{it}(0) & = & \theta_t' Z_i + \xi_t + \lambda_i' f_t + \varepsilon_{it} \\ Y_{it}(1) & = & Y_{it}^0 + \tau_{it} \end{cases}$$

- Let $W = (w_2, \ldots, w_{J+1})'$ with $w_j \geq 0$ and $w_2 + \cdots + w_{J+1} = 1$.
- Let $\bar{Y}_i^{K_1}, \ldots, \bar{Y}_i^{K_M}$ be $M > R$ linear functions of pre-intervention outcomes
- Suppose that we can choose W^* such that:

$$Z_1 = \sum_{j=2}^{J+1} w_j^* Z_j, \quad \bar{Y}_1^k = \sum_{j=2}^{J+1} w_j^* \bar{Y}_j^k, \; k \in \{K_1, \ldots, K_M\}$$

- When T_0 is large, an approximately unbiased estimator of τ_{1t} is:

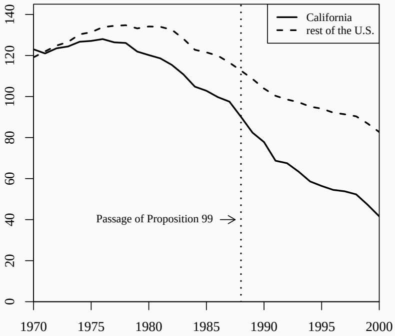
$$\widehat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}, \quad t \in \{T_0 + 1, \ldots, T\}$$

Implementation

- Let $X_1 = (Z_1, \bar{Y}_1^{K_1}, \ldots, \bar{Y}_1^{K_M})'$ be a $(k \times 1)$ vector of pre-intervention characteristics for the treated and X_0, a $(k \times J)$ matrix, for the controls.
- The vector W^* is chosen to minimize $\|X_1 - X_0 W\|$, subject to our weight constraints.
  - We consider $\|X_1 - X_0 W\|_V = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$, where V is some $(k \times k)$ symmetric and positive semidefinite matrix.
  - Various ways to choose V (subjective assessment of predictive power of X, regression, minimize MSPE, cross-validation, etc.).

Example: Proposition 99 on Cigarette Consumption

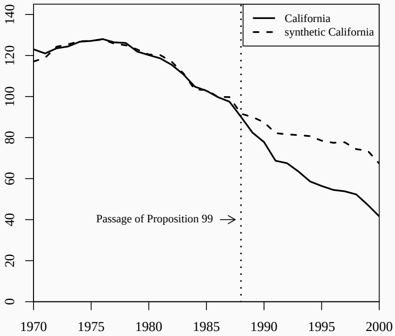
- In 1988, California first passed comprehensive tobacco control legislation (cigarette tax, media campaign etc.)
- Using 38 states that had never passed such programs as controls



Cigarette Consumption: CA and the Rest of the U.S.

Example: Proposition 99 on Cigarette Consumption

- In 1988, California first passed comprehensive tobacco control legislation (cigarette tax, media campaign etc.)
- Using 38 states that had never passed such programs as controls



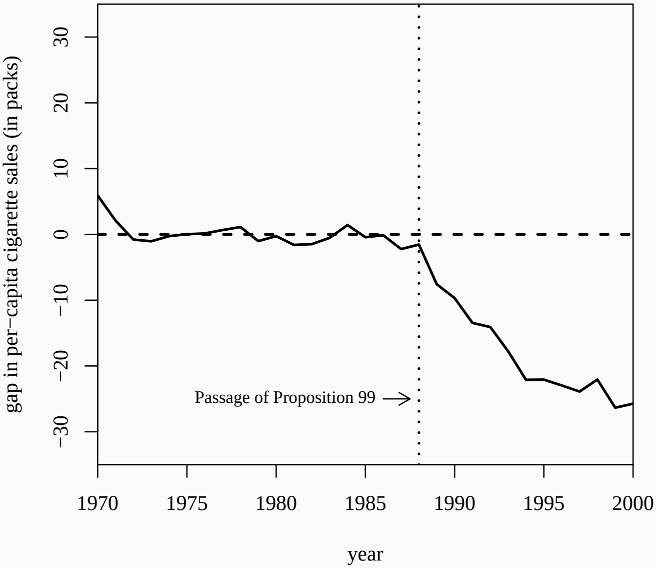
Cigarette Consumption: CA and Synthetic CA

Predictor Means: Actual vs. Synthetic California

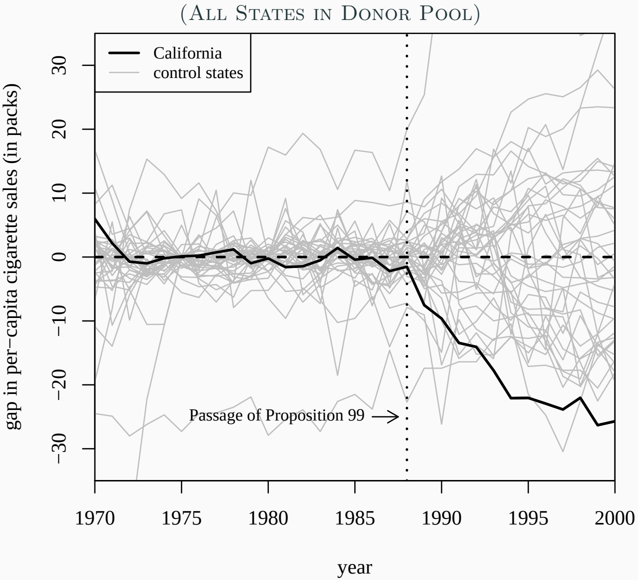
	California		Average of
Variables	Real	Synthetic	38 control states
Ln(GDP per capita)	10.08	9.86	9.86
Percent aged 15-24	17.40	17.40	17.29
Retail price	89.42	89.41	87.27
Beer consumption per capita	24.28	24.20	23.75
Cigarette sales per capita 1988	90.10	91.62	114.20
Cigarette sales per capita 1980	120.20	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).

Smoking Gap Between CA and Synthetic CA

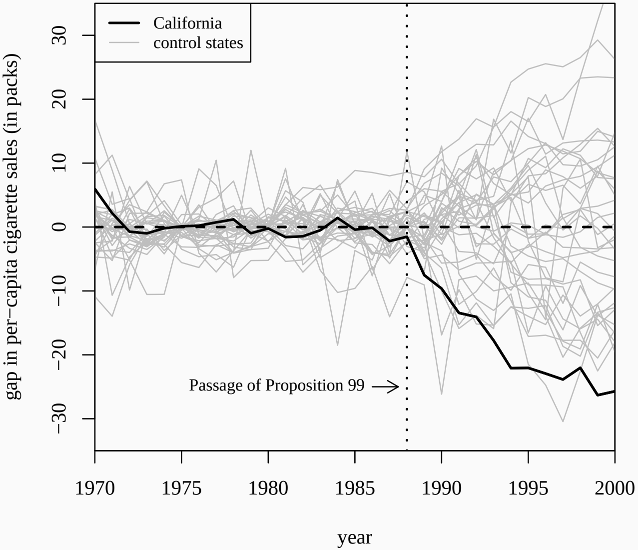


Smoking Gap for CA and 38 Control States



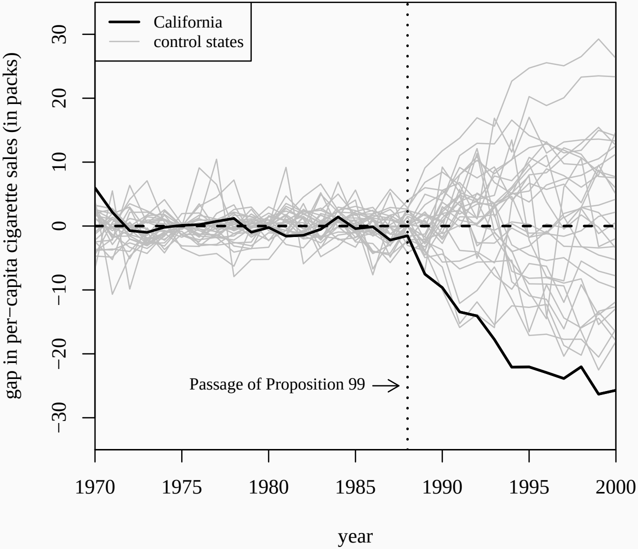
Smoking Gap for CA and 34 Control States

(Pre-Prop. 99 MSPE $\leq$ 20 Times Pre-Prop. 99 MSPE for CA)



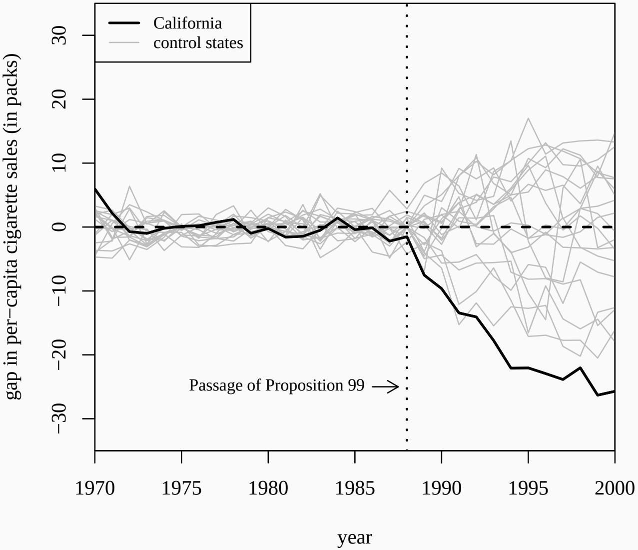
Smoking Gap for CA and 29 Control States

(Pre-Treatment MSPE $\leq$ 5 Times Pre-Treatment MSPE for CA)

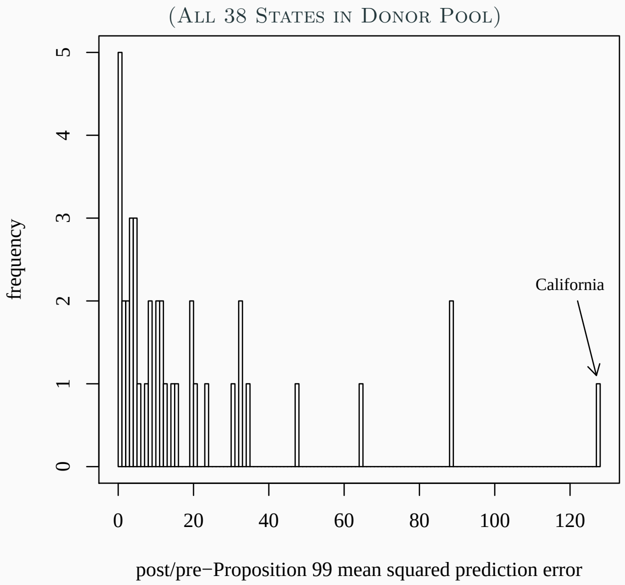


Smoking Gap for CA and 19 Control States

(Pre-Treatment MSPE $\leq$ 2 Times Pre-Treatment MSPE for CA)



Ratio Post-Treatment MSPE to Pre-Treatment MSPE



Limitations

- Algorithmic
  - Deal with one treated unit at a time
  - Deal with one outcome at a time
  - Slow to implement and sometimes difficult to find a solution
  - Allow too much user discretion, e.g. cherry-picking $\bar{Y}_i^k$ results in over-rejection (Ferman et al. 2017)
- Inference
  - Permutation inference and sensitivity analysis, (e.g. Hahn and Shi 2016; Firpo et al. 2017; Chernochukov 2017)
  - Inflated precision with nonstationary data (Cattaneo et al. 2019)
- Identification
  - My opinion: intrinsically, a method based on strict exogeneity (fixed timing)

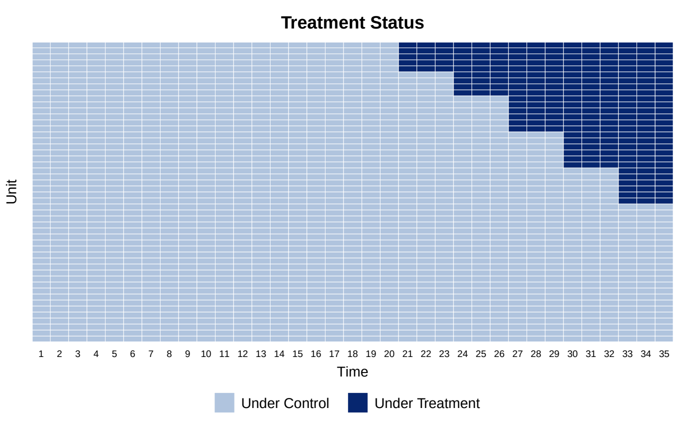
The Latent Factor Approach

Link with Synthetic Control

- Recall that ADH (2010) use a factor-augmented model to motivate the synthetic control method:
$$Y_{it}(0) = \theta_t' Z_i + \xi_t + \lambda_i' f_t + \varepsilon_{it}$$
- What if we actually estimate the model using observations under the control condition only?
- Xu (2017) imports the so-called interactive fixed-effect (IFE) model to a DID setting
$$Y_{it}(0) = X_{it}' \beta + \alpha_i + \xi_t + \lambda_i' f_t + \varepsilon_{it}$$
- Athey et al. (2021) extend it and introduce the matrix completion method
- Liu, Wang & Xu (2021) put these methods in a general framework — "the counterfactual estimators"
- No negative weighting!
- Limitations: needs large T and N; model dependency

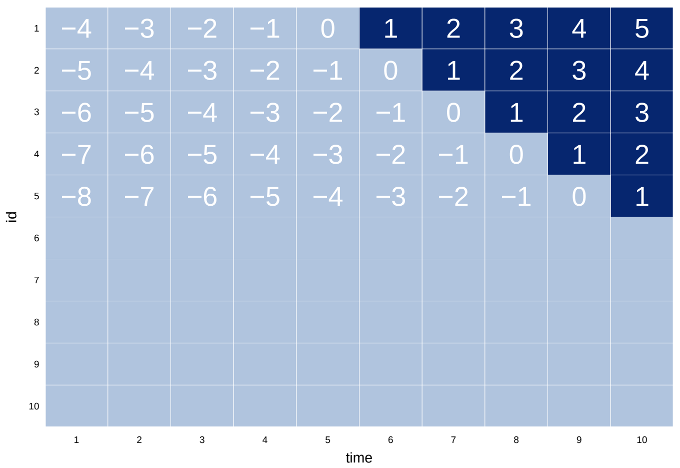
Basic Idea

- In a panel setting, treat $Y(1)$ as missing data
- Predict $Y(0)$ based on an outcome model
- (Use pre-treatment data for model selection)
- Estimate ATT by averaging differences between $Y(1)$ and $\hat{Y}(0)$

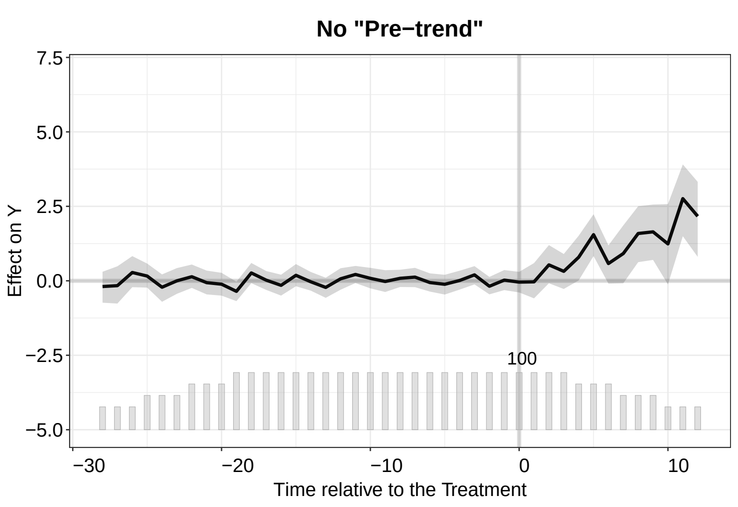


Basic Idea

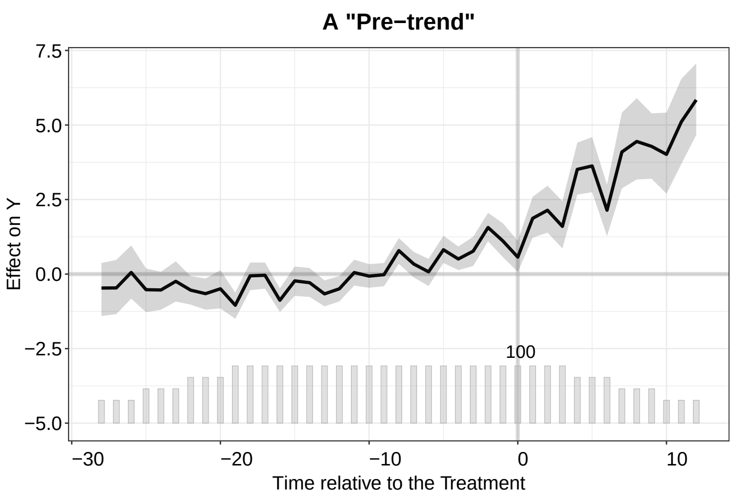
$$\widehat{ATT}_s = \hat{\mathbb{E}}[\hat{\tau}_{it} | D_{i,t-s} = 0, \underbrace{D_{i,t-s+1} = D_{i,t-s+2} = \cdots = D_{it} = 1}_{s \text{ periods}}, \forall i \in \mathcal{T}].$$



A New Plot for “Dynamic Treatment Effects”



A New Plot for “Dynamic Treatment Effects”



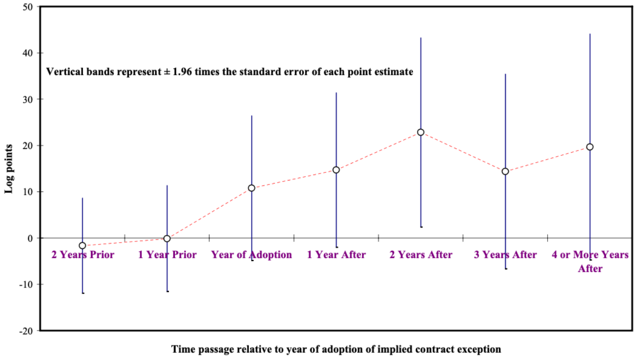


Figure 5.2.4: Estimated impact of state courts' adoption of an implied-contract exception to the employment-

Model-based Counterfactual Estimators

A model-based counterfactual estimator proceeds in the following steps:

- Step 1. Train the model using observations under the control condition ($D_{it} = 0$).
- Step 2. Predict the counterfactual outcome $\hat{Y}_{it}(0)$ for each observation under the treatment condition ($D_{it} = 1$) and obtain an estimate of the individual treatment effect: $\hat{\tau}_{it} = Y_{it} - \hat{Y}_{it}(0)$.
- Step 3. Generate estimates for the causal quantities of interest

$$ATT = \mathbb{E}[\tau_{it} | D_{it} = 1, \forall i \in \mathcal{T}, \forall t], \qquad \text{or}$$

$$ATT_s = \mathbb{E}[\tau_{it} | D_{i,t-s} = 0, \underbrace{D_{i,t-s+1} = D_{i,t-s+2} = \cdots = D_{it} = 1}_{s \text{ periods}}, \forall i \in \mathcal{T}].$$

Review of Three Estimators

We review three estimation strategies:

- FEct:
$$\hat{Y}_{it}(0) = X_{it}\hat{\beta} + \hat{\alpha}_i + \hat{\xi}_t$$
- IFEct (Gobillon&Magnac 2016; Xu 2017):
$$\hat{Y}_{it}(0) = X_{it}\hat{\beta} + \hat{\lambda}_i'\hat{F}_t$$
- Matrix Completion (MC) (Athey et al. 2018):
$$\hat{Y}_{it}(0) = X_{it}\hat{\beta} + \hat{L}_{it},$$
where matrix $\{L_{it}\}_{N\times T}$ is a lower-rank matrix approximation of $\{Y(0)\}_{N\times T}$ with missing values

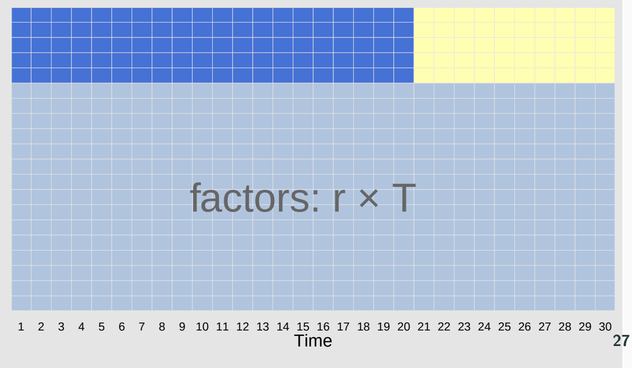
Remarks:

- DID is a special case of FEct
- Both IFEct and MC are estimated via iterative algorithms
- Cross-validation to choose the tuning parameter

Xu (2017) proposes a three-step approach based on a latent factor model:

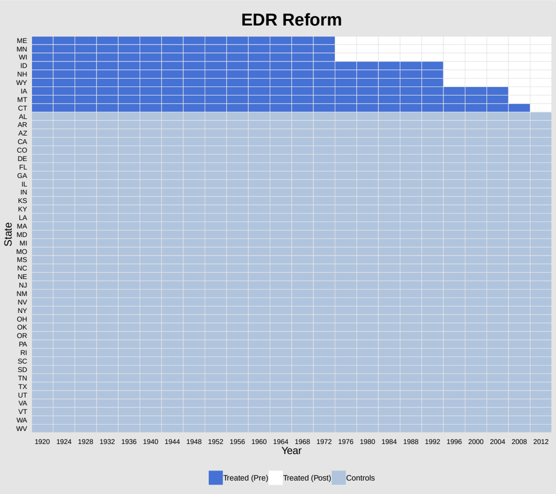
$$\begin{array}{llcll} \text{Control} & Y_{it}(0) & = & X_{it}'\beta + \alpha_i + \xi_t + \lambda_i' f_t + \varepsilon_{it} & \\ \text{Treated} & Y_{it}(0) & = & X_{it}'\beta + \alpha_i + \xi_t + \lambda_i' f_t + \varepsilon_{it} & \text{(pre)} \\ & Y_{it}(1) & = & X_{it}'\beta + \alpha_i + \xi_t + \lambda_i' f_t + \varepsilon_{it} + \tau_{it} & \text{(post)} \end{array}$$

1. Expectation-Maximization
   (Gobillon & Magnac 2016)

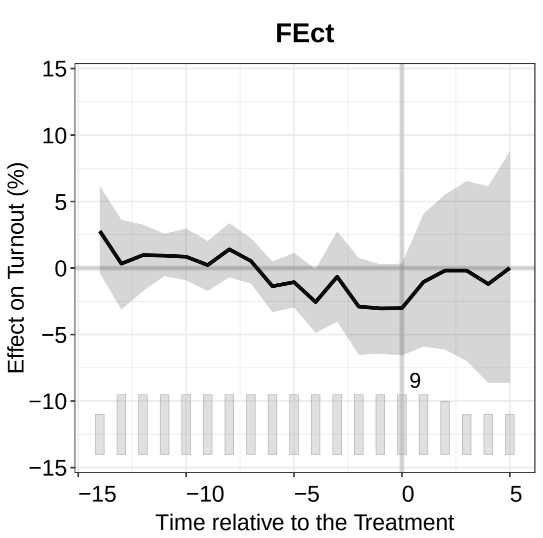


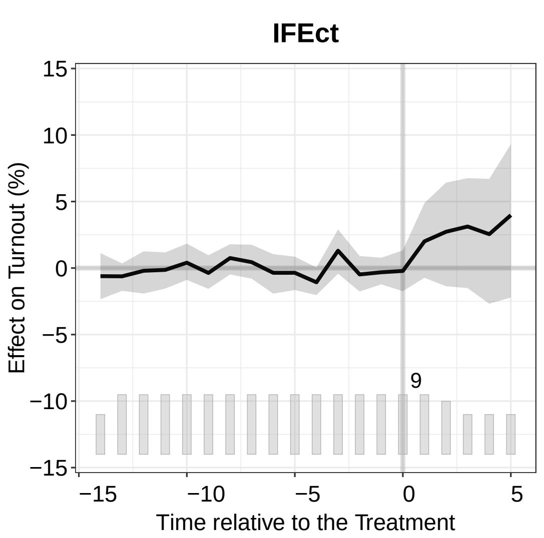
Election Day Registration (EDR) and Voter Turnout

Causal inference is a missing data problem.

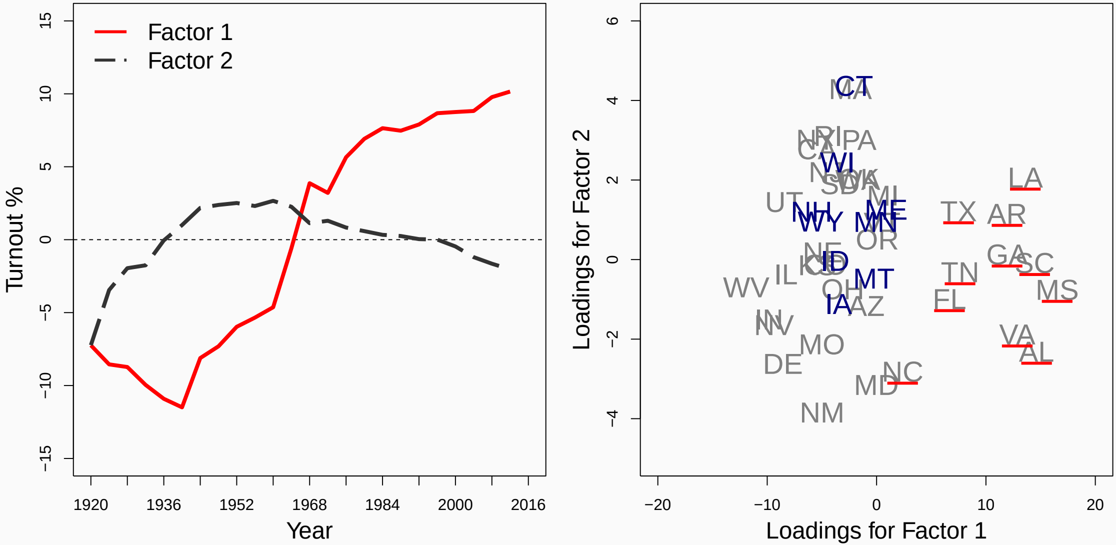


Main Results

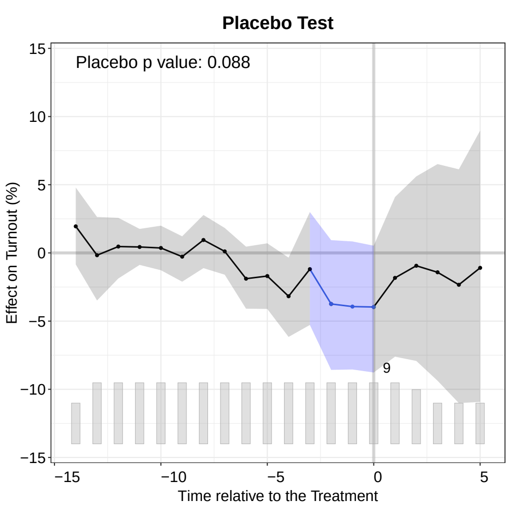


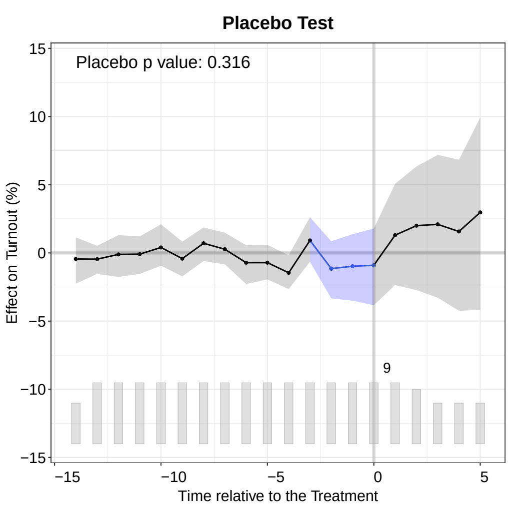


Factors and Factor Loadings



Placebo Test

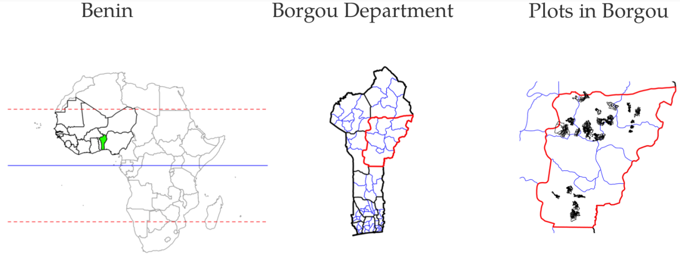




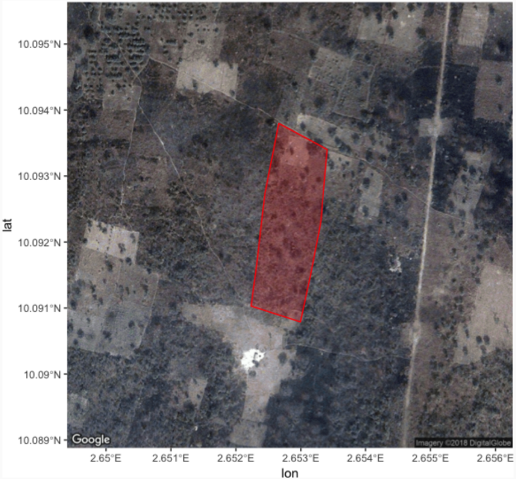
Example: Property Rights and Land Improvement

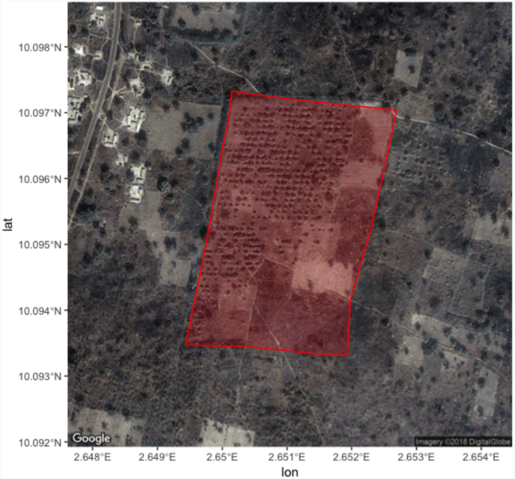
Sanford (2019)

- Does property rights lead to improved land quality?
- "Experiment": giving peasants in Borgou, Benin land titles
- Use satellite (remote sensing) data to measure land improvement, i.e., switch from annual crops to perennial crops (bushes and trees)
- Use IFEct to construct counterfactuals



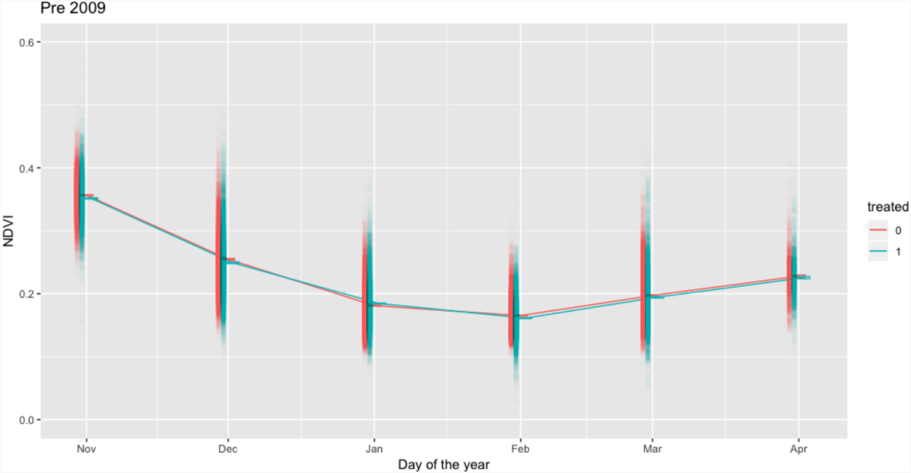
Original Satellite Images





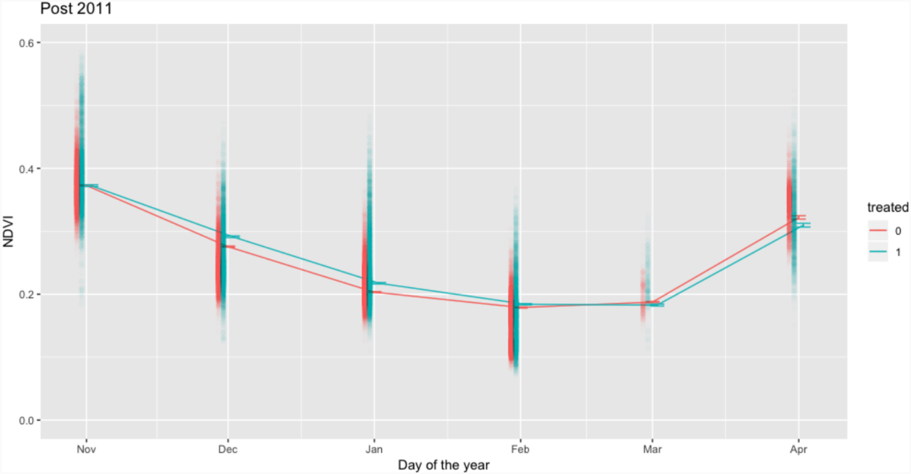
Treated and Counterfactual Averages

Pre-treatment Outcome



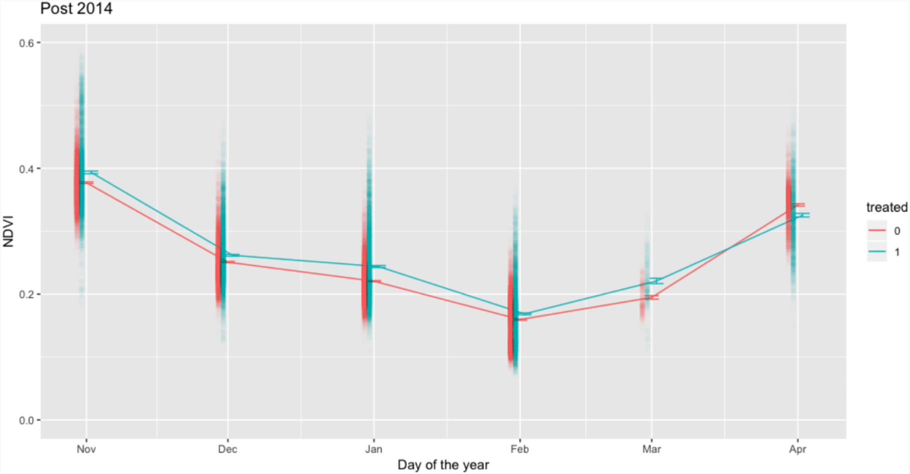
Treated and Counterfactual Averages

Post-treatment Outcome (1 Year)

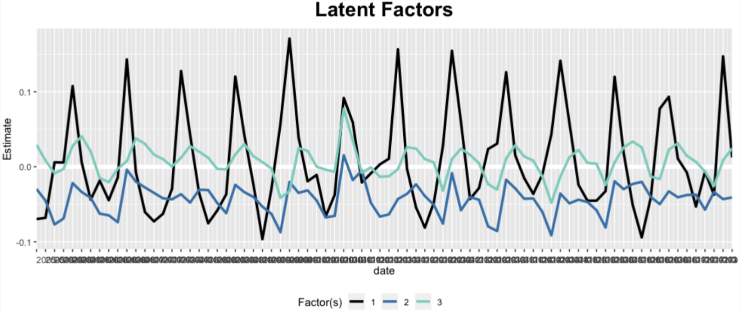


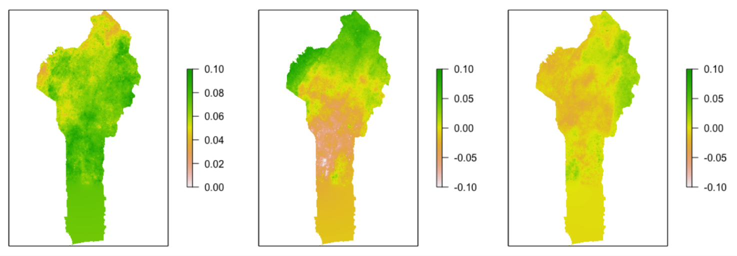
Treated and Counterfactual Averages

Post-treatment Outcome (4 Years)



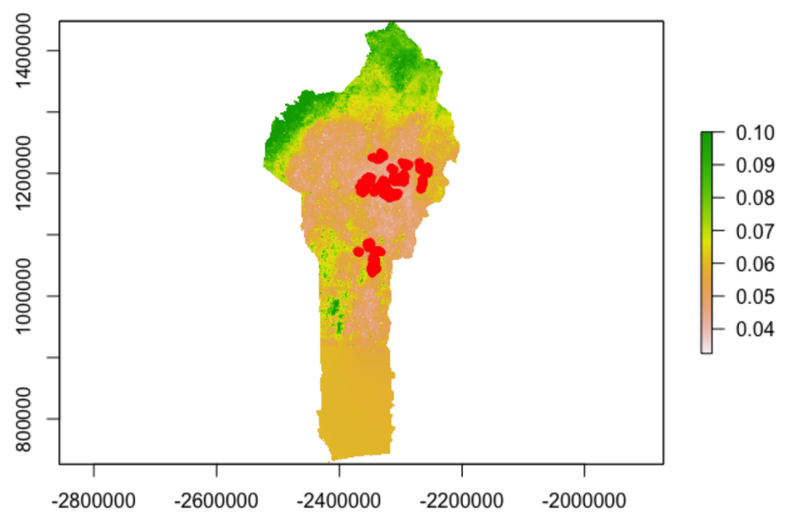
Factors and Loadings



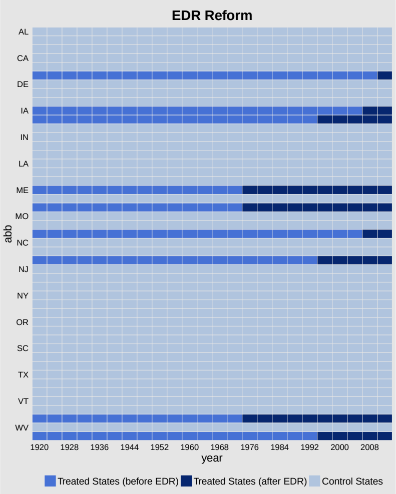


Geographic Distribution of Heavily-weighted Controls

Bigger number represents higher dissimilarity



Matrix Completion Methods



- Recall that our main goal is to predict treated counterfactuals
- Taking advantage of the matrix structure, matrix completion methods use non-treated data to achieve this goal
- The basic idea to find a lower-rank representation of the matrix to impute the "missing data"
- Xu (2017) is a special case of this approach

Matrix Completion Methods

- Recall in the baseline DID setup:

$$\mathbf{Y} = \begin{pmatrix} Y^0_{\mathcal{T},pre} & ?? \\ Y^0_{\mathcal{C},pre} & Y^0_{\mathcal{C},post} \end{pmatrix}$$

- Matrix completion (MC) methods attempt to find a lower-rank representation of **Y**, which we call **L**, that makes predictions of missing values in **Y**
- Athey et al. (2021) generalize Xu (2017) with different ways of constructing **L**
- Plus, missingness can be arbitrary $\rightarrow$ accommodate reversible treatments (note: strict exogeneity)

Matrix Completion Methods

- Mathematically,

$$Y_{it} = L_{it} + \alpha_i + \xi_t + X_{it}'\beta + \varepsilon_{it}$$

in which L_{it} is an element of $\mathbf{L}$, an $(N \times T)$ matrix

- We need regularization on $\mathbf{L}$ because of too many parameters:

$$\min_{\mathbf{L}} \frac{1}{\# Controls} \sum_{D_{it}=0} (Y_{it} - L_{it})^2 + \lambda_L \|\mathbf{L}\|_*$$

- The nuclear norm $\|.\|_*$ generally leads to a low-rank solution for $\mathbf{L}$

$$\|\mathbf{L}\|_* = \sum_{i=1}^{\min(N,T)} \sigma_i(\mathbf{L})$$

in which $\sigma_i(\mathbf{L})$ represents the i'th singular values of $\mathbf{L}$

IFEct vs. MC

- Singular value decomposition of L

$$\mathbf{L}_{N\times T} = \mathbf{S}_{N\times N}\boldsymbol{\Sigma}_{N\times T}\mathbf{R}_{T\times T}$$

- Difference in how $\boldsymbol{\Sigma}_{N\times T}$ is regularized

IFE	MC
best subset	nuclear norm

$$\begin{pmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \quad \begin{pmatrix} |\sigma_1 - \lambda_L|_+ & 0 & 0 & \cdots & 0 \\ 0 & |\sigma_2 - \lambda_L|_+ & 0 & \cdots & 0 \\ 0 & 0 & |\sigma_3 - \lambda_L|_+ & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ & & & & |\sigma_T - \lambda_L|_+ \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

in which $|a|_+ = \max(a, 0)$

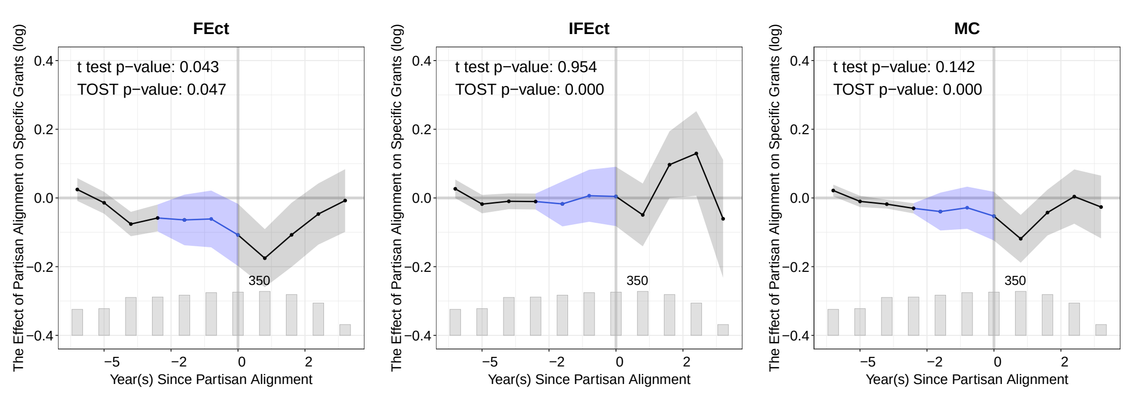
Inferential Methods and Diagnostics

- Non-parametric block bootstrap
  - sample with replacement across units
  - valid when N is large, $\frac{N_{tr}}{N}$ is fixed
- A permutation-based test for Sharp Nulls (Chernozhukov et al 2019)
  - e.g. $Y_{it}(1) = Y_{it}(0), \forall i \in \mathcal{T}, t > T_{0i}$
  - randomization over time (by blocks) instead of across units
  - valid if T is large, errors are stationary weekly dependent, and estimators are consistent or stable
  - exact if errors are i.i.d.
- Liu, Wang & Xu (2020) propose a set of diagnostic tests

Fouirnaies and Mutlu-Eren (2015)

Does partisan alignment brings special grant in UK?

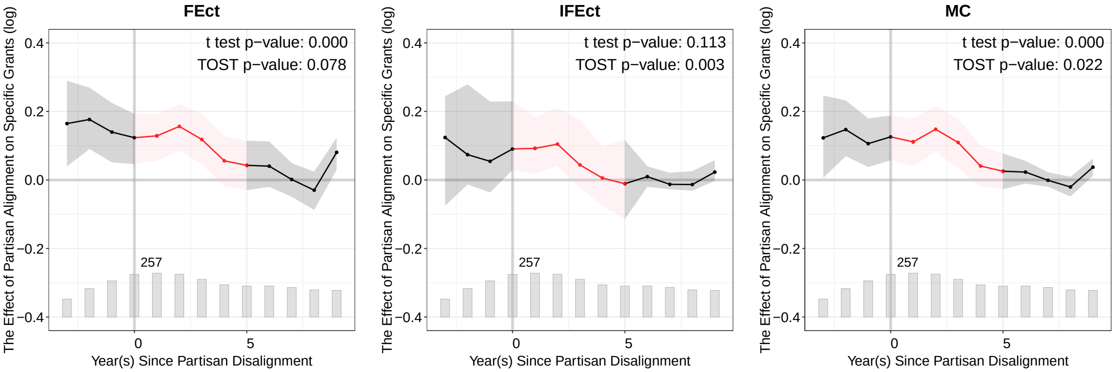
- Unit of analysis: 466 local councils from 1992 to 2012
- Treatment: Partisan alignment between locality and central government
- Outcome: Amount of special grant



Test for Carryover Effects

Does partisan alignment brings special grant in UK?

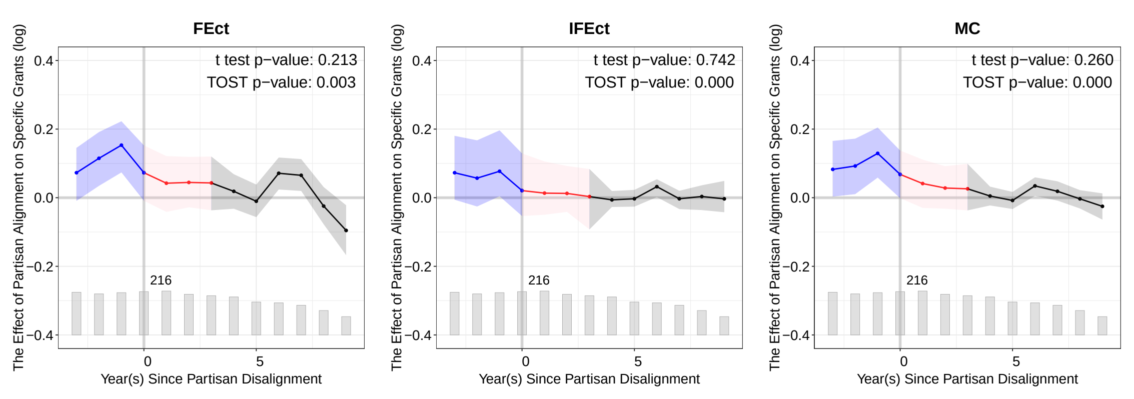
- Unit of analysis: 466 local councils from 1992 to 2012
- Treatment: Partisan alignment between locality and central government
- Outcome: Amount of special grant



Test for Carryover Effects (Removing 3 Periods)

Does partisan alignment brings special grant in UK?

- Unit of analysis: 466 local councils from 1992 to 2012
- Treatment: Partisan alignment between locality and central government
- Outcome: Amount of special grant



Doubly Robust Methods

Doubly Robust Methods

- So far, we've surveyed two group of methods: (1) those constructing balancing weights; (2) those modeling the conditional outcomes
- Combining the two approaches will likely produce doubly robust estimators
- Some methods we discussed, including semi-parametric DID, panel matching, trajectory balancing, are already doing a simple version of it (balancing plus regression)
- We review two new methods that formally adopt this idea
  - Augmented synthetic control (Ben-Michael et al 2018): modeling first
  - Synthetic DID (Arkhangelsky et al. 2019): weighting first

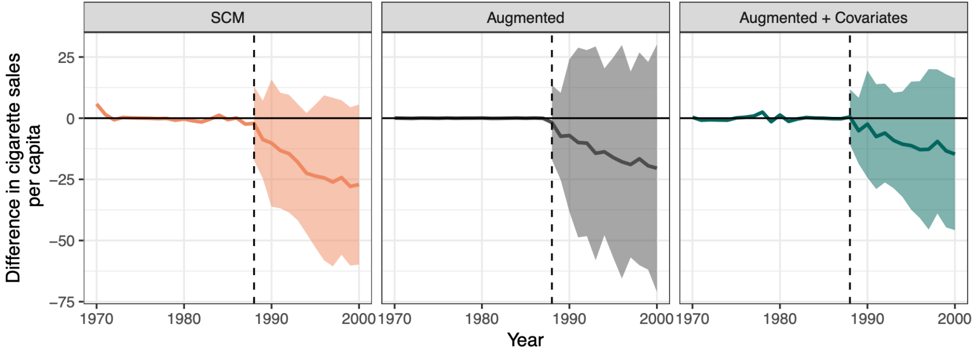
Augmented Synthetic Control (Ben-Michael et al 2017)

- Assuming unit 1 being treated ($D_1 = 1; D_{-1} = 0$), pretreatment covariates X_i
- **Basic Idea**
  1. Run an outcome model (e.g. Ridge, FEct, IFEct, MC, etc.) and obtain model fit $\hat{m}(X_i)$
  2. Balance on the residual averages, obtaining weights $\hat{\gamma}_i$ for the controls
  3. Treated average is constructed using:

$$\begin{aligned}\hat{Y}_1^{aug}(0) =& \underbrace{\sum_{i \in \mathcal{C}} \hat{\gamma}_i Y_i}_{SCM} + \underbrace{\hat{m}(X_1) - \sum_{i \in \mathcal{C}} \hat{\gamma}_i \hat{m}(X_i)}_{debias} \\ =& \hat{m}(X_1) + \sum_{i \in \mathcal{C}} \hat{\gamma}_i \left(Y_i - \hat{m}(X_i)\right)\end{aligned}$$

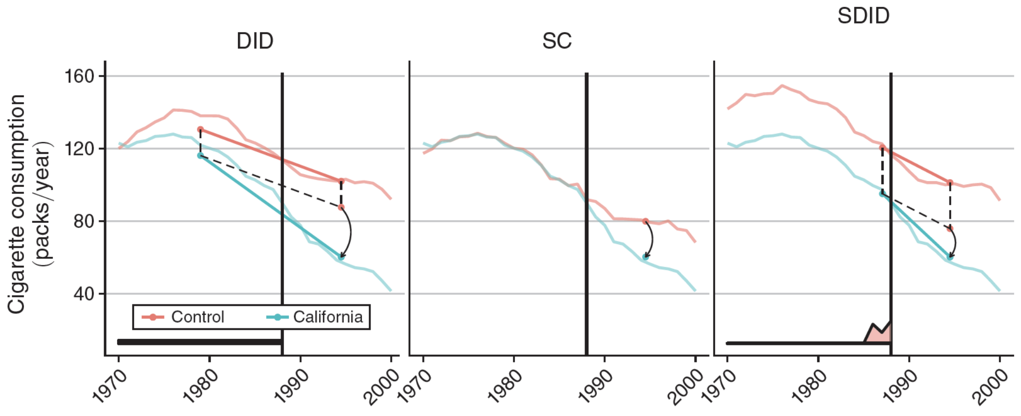
- The balancing weights take care of the remaining biases from the outcome model; the estimator is thus doubly robust
- Inference via jackknife

Revisiting California Prop 99



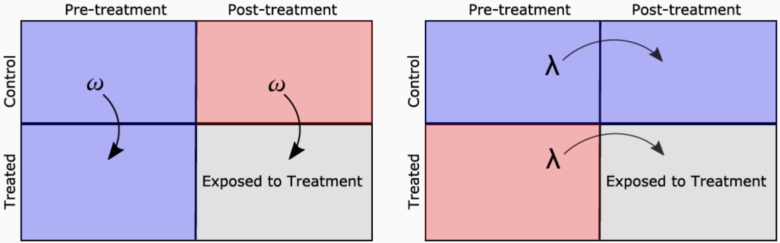
Synthetic DID (Arkhangelsky et al 2021)

- Key insight: weight not only control units, but also pretreatment periods (credit to David's slides)



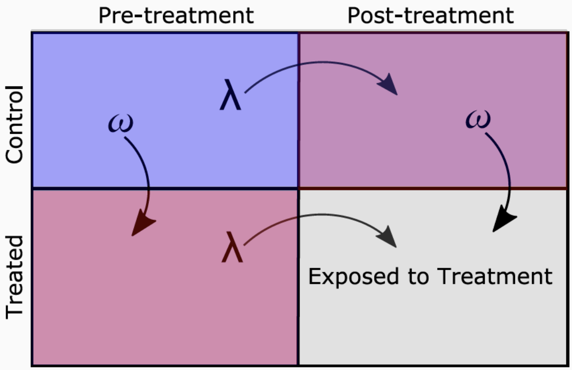
Unit and Time Weights

- SC:
  - Using pre-treatment data, we learn an average of controls that's predictive of California.
  - Assuming this relationship remain valid post-treatment, we use the same average of controls to impute treatment-free observations for California
- Forecasting:
  - Using controls, we learn an average of periods forecasting what we see post-treatment
  - Assuming this relationship remain valid for the treated, we use the same average of periods to impute treatment-free observations for California



Unit and Time Weights

- Synthetic Diff-in-Diff (SDID)
  - Do both synthetic control and forecasting and combine via diff-in-diff
  - Only one of these relationships has to remain valid (double robustness)
  - Our synthetic control can be parallel to California



SDID

1. Estimate unit weights $\hat{\omega}$ from pre-treatment data $\Rightarrow$ a synthetic control unit: $\hat{\omega}_0 + \hat{\omega}^T Y_{co,pre} \approx Y_{\overline{tr},pre}$.
2. Estimate time weights $\hat{\lambda}$ from control data $\Rightarrow$ a synthetic pre-treatment period: $\hat{\lambda}_0 + Y_{co,pre}\hat{\lambda} \approx Y_{co,\overline{post}}$.
3. Apply diff-in-diff to the resulting synthetic 2×2 panel by minimizing:

$$\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \mu - \alpha_i - \beta_t - X_{it}\gamma - D_{it}\tau)^2 \hat{\omega}_i \hat{\lambda}_t$$

	Synthetic Pre-Treatment	Average Post-treatment
Synthetic Control	$\hat{\omega}^T Y_{co,pre}\hat{\lambda}$	$\hat{\omega}^T Y_{co,\overline{post}}$
Average Treated	$Y_{\overline{tr},pre}\hat{\lambda}$	$Y_{\overline{tr},\overline{post}}$

* Inference via jackknife

Conclusions

Concluding Remarks

- The identification assumptions for DID are strong: functional form, no feedback, no spillover or general equilibrium effects
- TWFE models are often problematic: on top of DID assumptions, homogeneity (failure leads to negative weighting); limited carryover
- Counterfactual estimators, including the SCM, can be helpful but are not assumptions free
- Doubly robust methods have appealing statistical properties and will likely play large roles in empirical research

Practical Recommendations

- Plotting raw data, especially the distribution of treatment status, helps us see obvious problems
- Think harder on how the treatment is assigned; ask yourself: "what's the hypothetical experiment?"
- If you think feedback is weak, start from estimators under parallel trends (e.g., DID, DID_M, fect, AugSynth, SDID) and check "pre-trend"
- Diagnostics is important... Whichever method you use, conduct placebo tests to check if your identification assumptions are reasonable
- Inference is hard; cluster-bootstrap and jackknife (esp. when N_{tr} is small) are relatively safe choices

References

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). *Journal of the American Statistical Association*. June 1, 2010, 105(490): 493–505.
- Athey, Susan, and Guido Imbens. 2016. "The State of Applied Econometrics - Causality and Policy Evaluation." arXiv [stat.ME]. arXiv. http://arxiv.org/abs/1607.00699.
- Hahn, Jinyong, and Ruoyao Shi. 2017. "Synthetic Control and Inference." *Econometrics* 5 (4): 52.
- Firpo, Sergio, and Vitor Possebom. 2018. "Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets." *Journal of Causal Inference* 6 (2).
- Chernozhukov, Victor, Kaspar Wuthrich, and Yinchu Zhu. 2017. "An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls." arXiv [econ.EM]. arXiv. http://arxiv.org/abs/1712.09089.
- Ferman, B., C. Pinto, and V. Possebom. 2020. "Cherry Picking with Synthetic Controls." *Journal of Policy Analysis and Management*.
- Cattaneo, Matias D., Yingjie Feng, and Rocio Titiunik. 2019. "Prediction Intervals for Synthetic Control Methods." arXiv [stat.ME]. arXiv. http://arxiv.org/abs/1912.07120.
- Doudchenko, Nikolay and Guido Imbens (2016). "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis." Working Paper Stanford University.
- Hollingsworth, Alex, and Coady Wing. 2020. "Tactics for Design and Inference in Synthetic Control Studies: An Applied Example Using High-Dimensional Data." https://doi.org/10.2139/ssrn.3592088.
- Hsiao, Cheng, H. Steve Ching and Shui Ki Wan (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*, Vol. 27, Iss. 5, August 2012, pp. 705–740.
- Robbins, Michael W., Jessica Saunders, and Beau Kilmer. 2017. "A Framework for Synthetic Control Methods With High-Dimensional, Micro-Level Data: Evaluating a Neighborhood-Specific Crime Intervention." *Journal of the American Statistical Association* 112 (517): 109–26.

References

- Gobillon, Laurent, and Thierry Magnac (2016). "Regional Policy Evaluation: Interactive Fixed Effects and Synthetic Controls." Review of Economics and Statistics 98 (3): 535–51.
- Xu, Yiqing (2017). "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models" *Political Analysis*, Vol. 25, Iss. 1, January 2017, pp. 57-76.
- Athey, Susan, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. 2021. "Matrix Completion Methods for Causal Panel Data Models." *Journal of the American Statistical Association*, forthcoming. https://doi.org/10.3386/w25132.
- Liu, Licheng, Ye Wang, Yiqing Xu (2021). "A Practical Guide to Counterfacutal Estimators for Causal Inference with Time-Series Cross-Sectional Data." Working Paper, Stanford University.
- Ben-Michael Eli, Avi Feller, Jesse Rothstein (2018). "The Augmented Synthetic Control Method." Working Paper, UC Berkeley.
- Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens and Stefan Wager (2019). "Synthetic Difference In Differences." Working Paper, UC Berkeley.