Synthetic Difference-in-Differences

(Arkhangelsky et al, 2021)

Giryang Park

AERG

January 9, 2025

Contents

Introduction

Motivation

- Goal: Evaluate the effect of policy changes using panel data
- To obtain ATT, the key is to impute counterfactual for the treated

  $\rightarrow$ How to connect observed data to unobserved counterfactuals?

  $\rightarrow$ Two popular choices are Diff-in-Diff & Synthetic Control

Difference-in-Differences (DiD)	Many treated, Parallel trend
Synthetic Control (SC)	Single treated, Matching trend

Motivation

- Goal: Evaluate the effect of policy changes using panel data
- To obtain ATT, the key is to impute counterfactual for the treated

  $\rightarrow$ How to connect observed data to unobserved counterfactuals?

  $\rightarrow$ Two popular choices are Diff-in-Diff & Synthetic Control

Difference-in-Differences (DiD)	Many treated, Parallel trend
Synthetic Control (SC)	Single treated, Matching trend

✓ **Synthetic Difference-in-Differences (SDiD) integrates attractive features of DiD and SC, with competitive or even dominating performance**

Motivation: Comparison with DiD and SC

Basic Setup

- Balanced panel with N units and T time periods
- $N = N_{co} + N_{tr}$, $T = T_{pre} + T_{post}$
- Realized outcome Y_{it}, Binary treatment W_{it}
- Parameter of interest τ

Solve the following:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \left\{ \Sigma_{i,t} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \quad (1)$$

$$(\hat{\tau}^{did}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \left\{ \Sigma_{i,t} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \right\} \quad (2)$$

$$(\hat{\tau}^{sc}, \hat{\mu}, \hat{\beta}) = \underset{\tau,\mu,\beta}{\operatorname{argmin}} \left\{ \Sigma_{i,t} (Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sc} \right\} \quad (3)$$

Motivation: Comparison with DiD and SC

Basic Setup

- Balanced panel with N units and T time periods
- $N = N_{co} + N_{tr}$, $T = T_{pre} + T_{post}$
- Realized outcome Y_{it}, Binary treatment W_{it}
- Parameter of interest τ

Solve the following:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \left\{ \Sigma_{i,t} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \quad (1)$$

$$(\hat{\tau}^{did}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \left\{ \Sigma_{i,t} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \right\} \quad (2)$$

$$(\hat{\tau}^{sc}, \hat{\mu}, \hat{\beta}) = \underset{\tau,\mu,\beta}{\operatorname{argmin}} \left\{ \Sigma_{i,t} (Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sc} \right\} \quad (3)$$

Motivation: Comparison with DiD and SC

Basic Setup

- Balanced panel with N units and T time periods
- $N = N_{co} + N_{tr}$, $T = T_{pre} + T_{post}$
- Realized outcome Y_{it}, Binary treatment W_{it}
- Parameter of interest τ

Solve the following:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \{\Sigma_{i,t}(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid}\} \quad (1)$$

$$(\hat{\tau}^{did}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\operatorname{argmin}} \{\Sigma_{i,t}(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2\} \quad (2)$$

$$(\hat{\tau}^{sc}, \hat{\mu}, \hat{\beta}) = \underset{\tau,\mu,\beta}{\operatorname{argmin}} \{\Sigma_{i,t}(Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sc}\} \quad (3)$$

Key Contributions

1. New estimator reconciling vertical regression, horizontal regression, DiD regression approaches
2. Systematization of unit and time choice
3. Evaluation procedure using placebo studies based on nonuniform treatment assignment mechanism

Algorithm

Implementing SDiD

Algorithm 1—SDID

Data: $\mathbf{Y}, \mathbf{W}$
Result: Point estimate $\hat{\tau}^{sdid}$

1. Compute regularization parameter ζ using (5);
2. Compute unit weights $\hat{\omega}^{sdid}$ via (4);
3. Compute time weights $\hat{\lambda}^{sdid}$ via (6);
4. Compute the SDID estimator via the weighted DID regression

$$\left(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau\right)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\};$$

Implementing SDiD

Step 1) Compute regularization parameter ζ via

$$\zeta = (N_{tr} T_{post})^{1/4} \hat{\sigma}$$

Step 2) Compute unit weights $\hat{\omega}^{sdid}$ via

$$(\hat{\omega}_0, \hat{\omega}^{sdid}) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\operatorname{argmin}} \ \ell_{unit}(\omega_0, \omega) \ \text{ where}$$

$$\ell_{unit}(\omega_0, \omega) = \sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 + \zeta^2 T_{pre} \|\omega\|_2^2,$$

$$\Omega = \left\{ \omega \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} \omega_i = 1, \, \omega_i = N_{tr}^{-1} \text{ for all } i = N_{co}+1, \ldots, N \right\}$$

Implementing SDiD

Step 3) Compute time weights $\hat{\lambda}^{sdid}$ via

$$(\hat{\lambda}_0, \hat{\lambda}^{sdid}) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\operatorname{argmin}} \ \ell_{unit}(\lambda_0, \lambda) \ \text{ where}$$

$$\ell_{unit}(\lambda_0, \lambda) = \sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right)^2,$$

$$\Lambda = \left\{ \lambda \in \mathbb{R}_+^T : \sum_{t=1}^{T_{pre}} \lambda_t = 1, \ \lambda_t = T_{post}^{-1} \text{ for all } t = T_{pre}+1, \ldots, T \right\}$$

Step 4) Compute SDiD estimator via the weighted DiD regression (1)

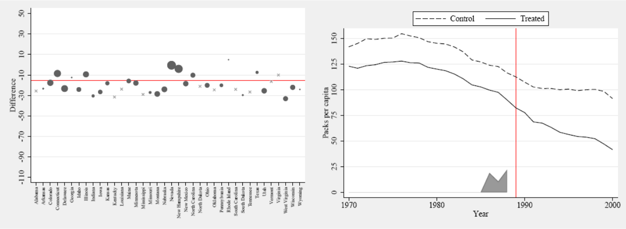
Implemented Results

California Smoking Cessation Program (Abadie, Diamond, and Hainmueller, 2010)

Synthetic Difference-in-Differences Estimator

packsperca~a	ATT	Std. Err.	t	P>|t|	[95% Conf.	Interval]
treated	-15.60383	9.53183	-1.64	0.102	-34.28588	3.07822

95% CIs and p-values are based on Large-Sample approximations.
Refer to Arkhangelsky et al., (2020) for theoretical derivations.



(a) Unit-Specific Weights (b) Outcome Trends and Time-Specific Weights

Figure 1: Proposition 99, example from Abadie et al. (2010); Arkhangelsky et al. (2021)

Implemented Results

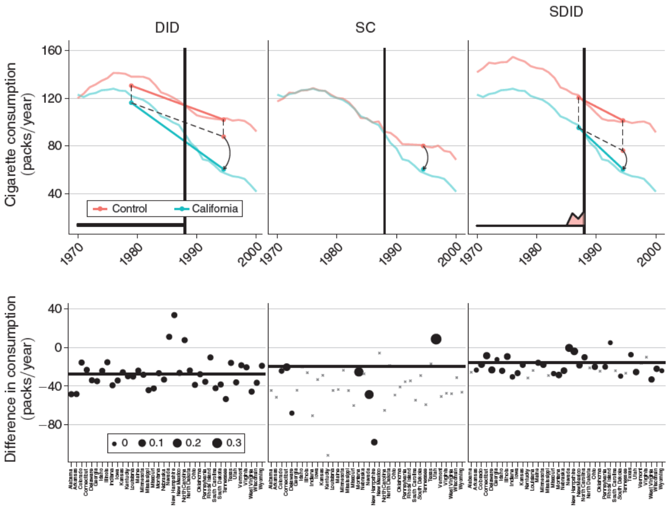


FIGURE 1. A COMPARISON BETWEEN DID, SC, AND SDID ESTIMATES FOR THE EFFECT OF CALIFORNIA PROPOSITION 99 ON PER-CAPITA ANNUAL CIGARETTE CONSUMPTION (IN PACKS/YEAR)

Formal Results

Latent Factor Model (Interactive Fixed Effects Model)

$$\mathsf{Y} = \mathsf{L} + \mathsf{W} \cdot \tau + \mathsf{E} \quad \text{where} \quad (W \cdot \tau)_{it} = W_{it}\tau_{it}$$

i.e., outcome is low rank matrix plus noise

- Systemtatic component $\mathsf{L} = \Gamma \Upsilon^T$
  : Γ is latent unit factor and Υ is latent time factor
- Treatment assignment W
- Error matrix E such that $\mathbb{E}(\mathsf{E} \mid \mathsf{W}, \mathsf{L}) = 0$
- Non-random assignment, that is, $\mathsf{W} \not\perp\!\!\!\perp \mathsf{L}$

Latent Factor Model (Interactive Fixed Effects Model)

$$\mathrm{Y} = \mathrm{L} + \mathrm{W} \cdot \tau + \mathrm{E} \quad \text{where} \quad (W \cdot \tau)_{it} = W_{it}\tau_{it}$$

i.e., outcome is low rank matrix plus noise

WANT: Recover ATT during the periods they were treated,

$$\tau = \frac{1}{N_{tr}\, T_{post}} \sum_{i=N_{co}+1}^{N} \sum_{t=T_{pre}+1}^{T} \boldsymbol{\tau}_{it} \tag{4}$$

Instead of directly estimating L, use <u>reweighting</u> instead.

$\rightarrow$ Balance out Γ so that $\Sigma_{tr}\, \hat{\omega}_i^{sc}\Gamma_i \approx \Sigma_{co}\hat{\omega}_i^{sc}\Gamma_i$ (or balance out Υ)

cf. Double robustness property

Weighted Double-Differencing Estimator

- Define weighted double-differencing estimator as equation (5)

$$\begin{aligned}\hat{\tau}(\omega, \lambda) = &(\omega_{tr}^T \mathsf{Y}_{tr,post}\lambda_{post} - \omega_{tr}^T \mathsf{Y}_{tr,pre}\lambda_{pre}) \\ &- (\omega_{co}^T \mathsf{Y}_{co,post}\lambda_{post} - \omega_{co}^T \mathsf{Y}_{co,pre}\lambda_{pre})\end{aligned} \tag{5}$$

- Noticing that τ can be recovered from the term $\omega_{tr}^T \mathsf{Y}_{tr,post}\lambda_{post}$, bias-noise decomposition proceeds as follows

$$\begin{aligned}&\hat{\tau}(\omega, \lambda) - \tau \\ &\quad = \underbrace{\omega_{tr}^T \mathsf{L}_{tr,post}\lambda_{post} - \omega_{tr}^T \mathsf{L}_{tr,pre}\lambda_{pre} - \omega_{co}^T \mathsf{L}_{co,post}\lambda_{post} + \omega_{co}^T \mathsf{L}_{co,pre}\lambda_{pre}}_{\text{bias } B(\omega,\lambda)} \\ &\quad + \underbrace{\omega_{tr}^T \mathsf{E}_{tr,post}\lambda_{post} - \omega_{tr}^T \mathsf{E}_{tr,pre}\lambda_{pre} - \omega_{co}^T \mathsf{E}_{co,post}\lambda_{post} + \omega_{co}^T \mathsf{E}_{co,pre}\lambda_{pre}}_{\text{noise } \varepsilon(\omega,\lambda)}\end{aligned} \tag{6}$$

Oracle Weights and Error Decomposition

- Oracle weights $\tilde{\omega}, \tilde{\lambda}$ refer to 'ideal' weights which optimize population criterion (7) under assumption of perfect knowledge on DGP

$$(\tilde{\omega}_0, \tilde{\omega}) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\operatorname{argmin}} \mathbb{E}[\ell_{unit}(\omega_0, \omega)], \quad (\tilde{\lambda}_0, \tilde{\lambda}) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\operatorname{argmin}} \mathbb{E}[\ell_{time}(\lambda_0, \lambda)] \quad (7)$$

- Deterministic oracle weights have desirable properties; estimator using oracle weights predict counterfactuals accurately and satisfy CLT around its expectation

Oracle Weights and Error Decomposition

- Further decompose with oracle weights

$$
\begin{aligned}
\hat{\tau}^{sdid} - \tau &= \hat{\tau}(\hat{\omega}, \hat{\lambda}) - \tau \\
&= \underbrace{\varepsilon(\tilde{\omega}, \tilde{\lambda})}_{\text{oracle noise}} + \underbrace{B(\tilde{\omega}, \tilde{\lambda})}_{\text{oracle confounding bias}} + \underbrace{\hat{\tau}(\hat{\omega}, \hat{\lambda}) - \hat{\tau}(\tilde{\omega}, \tilde{\lambda})}_{\text{deviation from oracle}}
\end{aligned} \tag{8}
$$

- Oracle noise is small when
  1) weights are not too concentrated
  2) either N_{tr} or T_{post} is large, but both still small relative to N_{co}, T_{pre}
- Oracle confounding bias is small when pre-exposure trend fits well
- Deviation from oracle is small when above conditions are satisfied

Consistency and Asymptotic Gaussianity

Under some assumptions and regularity conditions including

- Rank of L small relative to $\sqrt{\min(N_{co}, T_{pre})}$
- Units are exchangeable, yet time periods are not exchangeable
- Rows of E are i.i.d and gaussian,

$$\hat{\tau}^{sdid} - \tau = \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} \varepsilon_{it} - \mathrm{E}_{i,pre}\psi \right) + o_p((N_{tr} T_{post})^{-1/2}) \quad (7)$$

and consequently,

$$\left(\hat{\tau}^{sdid} - \tau\right) / V_{\tau}^{1/2} \Rightarrow \mathcal{N}(0,1) \quad (8)$$

holds.

Properties of SDiD Estimator

- Two facets of double robustness property

Unit Weight and Time Weight

If either of unit balancing or time balancing is effective, bias from L is approximately removed.

Model and Weights

1. Under general factor model, with good weights, SDiD is consistent.
2. Under DiD model, even with bad weights, SDiD is consistent.

- SDiD estimator has good bias properties.

Large Sample Inference

In order to use the following asymptotics

$$\tau \in \hat{\tau}^{sdid} \pm z_{\alpha/2}\sqrt{\hat{V}_{\tau}}, \tag{9}$$

need to construct consistent estimator for asymptotic variance V_{τ}.

$\rightarrow$ Three options

1. Clustered Bootstrap
   - Resample and utilize distribution of estimated treatment effect
2. Jackknife
   - Compute leave-one-out version of estimated treatment effect
3. Placebo Approach
   - Suppose some control units were exposed to the treatment

Placebo Studies

Data Generating Process

- Revisits benchmark placebo study on DiD estimator (Bertrand, Duflo, and Mullainathan, 2004)
  - Key idea: Randomly assign a subset of states to a placebo treatment

    $\Rightarrow$ True treatment effect of zero, by construction
  - Modifications
    - Introduction of non-uniform treatment assignment mechanism
    - Simulated values for outcomes via model estimated on CPS data

Data Generating Process

Outcome Model

- Outcome: Wages for women with positive wages (state/year aggregate)
- Fit L and Σ where $\mathrm{E}_i \sim \mathcal{N}(0, \Sigma)$ using CPS data

$$\mathrm{L} := \underset{L:\mathrm{rank}(L)=4}{\mathrm{argmin}} \ \Sigma_{i,t}(Y_{it}^* - L_{it})^2, \quad \Sigma : \mathrm{AR(2)\ w/\ residuals} \tag{10}$$

Assignment Model

- Use binary exposure indicators related to state policy: $W_{it} = D_i 1_{t > T_{pre}}$
- Reflects actual differences on important state-wise economic conditions
  ex) minimum wage laws, abortion rights, gun control laws

$$\begin{gathered} D_i \mid \mathrm{E}_i, \alpha_i, \mathrm{M}_i \sim \mathrm{Bern}(\pi_i) \\ \pi_i = \pi(\alpha_i, \mathrm{M}_i; \phi) = \frac{\exp(\phi_\alpha \alpha_i + \phi_M \mathrm{M}_i)}{1 + \exp(\phi_\alpha \alpha_i + \phi_M \mathrm{M}_i)} \end{gathered} \tag{11}$$

Simulation Results

Table 2

	RMSE					Bias				
	SDID	SC	DID	MC	DIFP	SDID	SC	DID	MC	DIFP
1. Baseline	**0.28**	0.37	0.49	0.35	0.32	0.10	0.20	0.21	0.15	0.07
Outcome model										
2. No corr	**0.28**	0.38	0.49	0.35	0.32	0.10	0.20	0.21	0.15	0.07
3. No **M**	0.16	0.18	**0.14**	**0.14**	0.16	0.01	0.04	0.01	0.01	0.01
3. No **F**	0.28	**0.23**	0.49	0.35	0.32	0.10	0.04	0.21	0.15	0.07
4. Only noise	0.16	**0.14**	**0.14**	**0.14**	0.16	0.01	0.01	0.01	0.01	0.01
5. No noise	0.06	0.17	0.47	**0.04**	0.11	0.05	0.04	0.20	0.00	0.01
Assignment process										
6. Gun law	**0.26**	0.27	0.47	0.36	0.30	0.08	−0.03	0.15	0.15	0.09
7. Abortion	**0.23**	0.31	0.45	0.31	0.27	0.04	0.16	0.03	0.02	0.01
8. Random	**0.24**	0.25	0.44	0.31	0.27	0.01	−0.01	0.02	0.01	−0.00
Outcome variable										
9. Hours	1.90	2.03	2.06	**1.85**	1.97	1.12	−0.49	0.85	1.00	1.00
10. U-rate	**2.25**	2.31	3.91	2.96	2.30	1.77	1.73	3.60	2.63	1.69
Assignment block size										
11. $T_{post} = 1$	**0.50**	0.59	0.70	0.51	0.54	0.20	0.17	0.38	0.21	0.12
12. $N_{tr} = 1$	**0.63**	0.73	1.26	0.81	0.83	0.03	0.15	0.11	0.05	−0.02
13. $T_{post} = N_{tr} = 1$	1.12	1.24	1.52	**1.07**	1.16	0.14	0.24	0.33	0.16	0.11

→ Particularly well-performing when treatment assignment is not uniformly random

Simulation Results

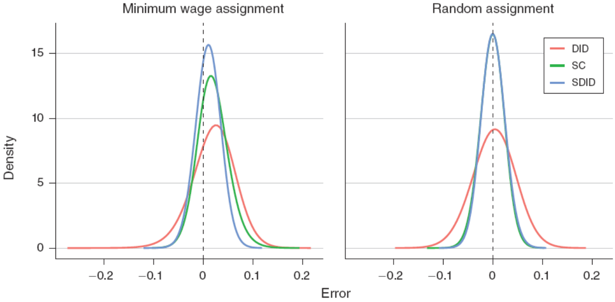


FIGURE 2. DISTRIBUTION OF THE ERRORS OF SDID, SC, AND DID IN THE SETTING OF THE "BASELINE" (I.E., WITH MINIMUM WAGE) AND RANDOM ASSIGNMENT ROWS OF TABLE 2

$*$ Refer to the paper for another simulation study using Penn World Table

Application

Code Implementation

- With **Stata**,
  sdid Y S T D [if] [in], vce(method) seed(#) reps(#) covariates(varlist [, method]) ...
  For more info, visit github repository `sdid`.

- With **R**,
  setup = panel.matrices(data)
  tau.hat = synthdid_estimate(data$Y, data$N0, data$T0)
  For more info, visit github webpage `synthdid`.

- With **Python**, package `synthdid` and so on ...

Application: Abadie, Diamond & Hainmueller(2010)

- Estimate the effect of increased cigarette tax on smoking in California
- State level panel data with 39 states, year 1970-2000
- 1 treated and 28 control, 19 pre-exposure and 12 post-exposure periods
- Y_{it}: per capita cigarette sales

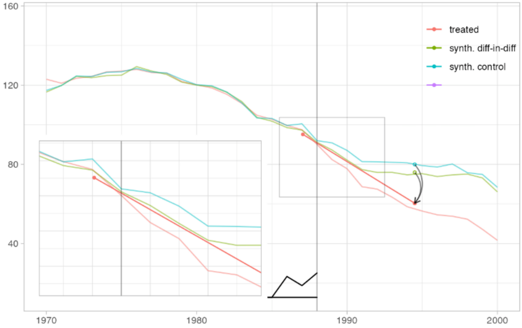


Figure: Source = Github Webpage R `synthdid`

Appendix

Regularity Conditions

- A1. The rows E_i of noise matrix are independent and identically distributed Gaussian vectors and the eigenvalues of its covariance matrix Σ are bounded and bounded away from zero.
- A2. A sequence of populations should satisfy
  (i) $N_{tr}T_{post}$ goes to infinity, and both N_{co} and T_{pre} go to infinity,
  (ii) T_{pre}/N_{co} is bounded and bounded away from zero,
  (iii) $N_{co}/(N_{tr}T_{post}\max(N_{tr}, T_{post})\log^2(N_{co})) \to \infty$.
- A3. Letting $\sigma_1(\Gamma), \sigma_2(\Gamma), \ldots$ denote singular values of the matrix Γ in decreasing order and R the largest integer less than $\sqrt{\min(T_{pre}N_{co})}$, $\sigma_R(\mathsf{L}_{co,pre})/R = o(\min\{(N_{tr}\log(N_{co}))^{-1/2}, (T_{post}\log(T_{pre}))^{-1/2}\})$.

Regularity Conditions

- A4-1. The oracle unit weights $\tilde{\omega}$ satisfy

$$||\tilde{\omega}_{co}||_2 = o([(N_{tr}T_{post})\log(N_{co})]^{-1/2})$$

and

$$\begin{aligned}&||\tilde{\omega}_0 + \tilde{\omega}_{co}^T \mathrm{L}_{co,pre} - \tilde{\omega}_{tr}^T \mathrm{L}_{tr,pre}||_2\\ &= o(N_{co}^{1/4}(N_{tr}T_{post}\max(N_{co}, T_{post}))^{-1/4}\log^{-1/2}(N_{co})).\end{aligned}$$

- A4-2. The oracle time weights $\tilde{\lambda}$ satisfy

$$||\tilde{\lambda}_{pre} - \psi||_2 = o([(N_{tr}T_{post})\log(N_{co})]^{-1/2})$$

and

$$||\tilde{\lambda}_0 + \mathrm{L}_{co,pre}\tilde{\lambda}_{pre} - \mathrm{L}_{co,post}\tilde{\lambda}_{post}||_2 = o(N_{co}^{1/4}(N_{tr}T_{post})^{-1/8}).$$

- A4-3. The oracle weights jointly satisfy

$$B(\tilde{\omega}, \tilde{\lambda}) = o((N_{tr}T_{post})^{-1/2}).$$

Staggered Assignment

- Block assignment $W_1 =$

$$\begin{array}{cc} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

- Staggered assignment $W_2 =$

$$\begin{array}{cc} & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

Staggered Assignment

- Split W_2 to make two distinct samples and fit estimators using block SDiD. Then calculate property weighted average of two estimators.

$$
\begin{array}{cc}
 & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} &
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}
\end{array}
\rightarrow
\left[
\begin{array}{cccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
4 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
\hdashline
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
5 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
6 & 0 & 0 & 1 & 1 & 1 & 1 & 1
\end{array}
\right]
$$

Reference

- Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. 2021. "Synthetic Difference-in-Differences." American Economic Review, 111 (12): 4088–4118.
- Damian Clarke, Daniel Pailañir, Susan Athey, Guido Imbens. 2023. "Synthetic Difference-in-Differences Estimation." IZA Discussion Papers, No.15907.
- Winter 2021 Online Causal Inference Seminar. 2021. "Susan Athey: Synthetic Difference in Differences." Video Link.