A Design-Based Perspective on Synthetic Control Methods

Lea Bottmer
lbottmer@stanford.edu

Guido W. Imbens
imbens@stanford.edu

Jann Spiess
jspiess@stanford.edu

Merrill J. Warnick
mwarnick@stanford.edu

Stanford University

October 7, 2021

Abstract

Since their introduction in Abadie and Gardeazabal (2003), Synthetic Control (SC) methods have quickly become one of the leading methods for estimating causal effects in observational studies in settings with panel data. Formal discussions often motivate SC methods by the assumption that the potential outcomes were generated by a factor model. Here we study SC methods from a design-based perspective, assuming a model for the selection of the treated unit(s) and period(s). We show that the standard SC estimator is generally biased under random assignment. We propose a Modified Unbiased Synthetic Control (MUSC) estimator that guarantees unbiasedness under random assignment and derive its exact, randomization-based, finite-sample variance. We also propose an unbiased estimator for this variance. We document in settings with real data that under random assignment, SC-type estimators can have root mean-squared errors that are substantially lower than that of other common estimators. We show that such an improvement is weakly guaranteed if the treated period is similar to the other periods, for example, if the treated period was randomly selected.

Keywords: Randomization, Panel Data, Causal Effects, Inference

Lea Bottmer is a PhD Candidate, Department of Economics, Stanford University, Stanford, CA 94305. Guido Imbens is The Applied Econometrics Professor and Professor of Economics, Graduate School of Business and Department of Economics, Stanford University, Stanford, CA 94305. Jann Spiess is Assistant Professor of Operations, Information and Technology, Graduate School of Business, Stanford University, Stanford, CA, 94305. Merrill Warnick is a PhD Candidate, Department of Economics, Stanford University, Stanford, CA, 94305. This research was generously supported by ONR grant N00014-17-1-2131. The authors would like to thank Alberto Abadie, Kirill Borusyak, Jonathan Roth, and Jeffrey Wooldridge for helpful comments and general discussions on the topic of this paper.

1 INTRODUCTION

Synthetic Control (SC) methods for estimating causal effects have become very popular in empirical work in the social sciences since their introduction by Alberto Abadie and coauthors (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015). Typically, the properties of the SC estimator are studied under model-based assumptions about the distribution of the potential outcomes in the absence of the intervention. For example, a common approach is to assume that the potential outcomes follow a factor model with noise. In this article, we take a design-based approach to synthetic control methods where instead of making assumptions about the distribution of the potential outcomes, we make assumptions about the assignment of the unit/time-period pairs to treatment. In particular, we consider the case with the treated unit(s) selected at random from a set of units. We find that in this setting the original SC estimator is generally biased. We propose a modification of the SC estimator, labeled the Modified Unbiased Synthetic Control (MUSC) estimator, which is unbiased under random assignment of the treatment. We also derive the exact variance of this estimator, and propose an unbiased estimator for this variance.

Studying the properties of SC-type estimators under design-based assumptions serves two distinct purposes. First, it suggests an important role for SC methods in the analysis of data from randomized experiments. Second, it leads to new insights into the properties of SC methods in observational studies. We show that SC methods are particularly valuable in experimental settings with relatively few units and multiple lagged outcomes, in the presence of substantial correlation over time and across units. In such experiments, where our design assumptions hold by definition, SC-type methods (including both the original SC estimator and our proposed MUSC estimator) can have substantially better root-mean-squared-error (RMSE) properties than the standard difference-in-means (DiM) estimator, with this improvement guaranteed under time randomization and a large number of time periods. However, it may be important to maintain the guaranteed unbiasedness that the DiM estimator enjoys under randomization. The proposed MUSC estimator does so, and combines the typical improvement in terms of RMSE with unbiasedness under randomization. We show that the proposed MUSC estimator extends naturally to cases with non-constant propensity scores and multiple treated units. Beyond the analysis of existing data, these results can be relevant for choosing assignment probabilities when planning a randomized experiment.

To illustrate the benefits of the MUSC estimator in experimental settings, we simulate an experiment based on data on average log wages observed across 10 states over 40 years. We randomly select one state to be treated in the last period, and compare (*i*) the DiM, (*ii*) the standard SC estimator, (*iii*) our proposed MUSC estimator, and (*iv*) DiD (difference-in-differences). There are three key findings, partially reported in Table 1 and expanded on in Section 5. First, the difference-in-means, the difference-in-differences, and the new MUSC estimator are unbiased by construction whereas the SC estimator is biased. Although in this example the bias of the SC estimator is modest, the bias need not be small in general. Second, the RMSE is substantially lower for the SC and MUSC estimators relative to the RMSE of the difference-in-means estimator, with the DiD estimator in the middle. This shows that there can be considerable gains to using SC methods in randomized experiments, without a need to give up the unbiasedness guaranteed by the randomization. Third, the proposed variance estimator is accurate in this setting for all four estimators.

Table 1: Simulation Experiment Based on CPS Average Log Wage by State and Year

	DiM	SC	MUSC	DiD
Bias	0	−0.007	0	0
RMSE	0.105	0.051	0.048	0.060
Average standard error	0.105	0.051	0.048	0.060

DiM: Difference in Means estimator; SC: Synthetic Control estimator of Abadie et al. (2010); MUSC: Modified Unbiased Synthetic Control estimator; DiD: Difference-in-Differences estimator.

The second contribution of the article concerns insights into synthetic-control methods in observational studies. These insights fall into four categories. First, we propose a new estimator (the MUSC estimator) for those settings that comes with additional robustness guarantees relative to the previously proposed SC estimators. Second, we develop new approaches to inference. Inference has been a challenge in the SC literature, especially in settings with a single treated unit and only a modest number of control units and time periods. We show that the placebo variance estimator, which is often used in that case and implicitly assumes that treatment assignment is exchangeable to some degree, can be biased under the random assignment assumption, while our proposed variance estimator is unbiased in finite samples, even with a single treated unit. Third, the design perspective highlights the importance of the choice of estimand. We show how the choice of estimand relates to the assumptions and inferential methods. Fourth, we show that the criterion for choosing the weights has some optimality conditions under random assignment of the treated period.

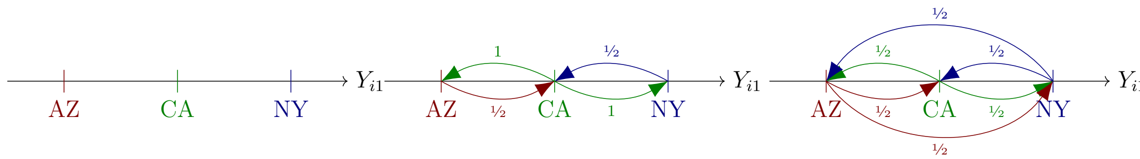
Although in observational studies random assignment is not guaranteed, we argue that it is still a natural starting point for many analyses, often after some adjustment for observed covariates. In particular, many of the analyses of SC methods explicitly or implicitly refer to units being comparable. This includes the placebo analyses used to test hypotheses (Abadie et al., 2010; Firpo and Possebom, 2018). In addition, many applications informally make reference to such assumptions as justification for the methods (Coffman and Noy, 2012; Cavallo et al., 2013; Liu, 2015).

The article builds on the general SC literature started by Abadie and Gardeazabal (2003); Abadie et al. (2010, 2015). See Abadie (2019) for a recent survey. It specifically contributes to the literature proposing new estimators for this setting, including Doudchenko and Imbens (2016); Abadie and L'Hour (2017); Ferman and Pinto (2017); Arkhangelsky et al. (2019); Li (2020); Ben-Michael et al. (2020). The current paper also contributes to the literature on inference for SC estimators, which includes Abadie et al. (2010); Doudchenko and Imbens (2016); Ferman and Pinto (2017); Hahn and Shi (2016); Lei and Candès (2020); Chernozhukov et al. (2017). We also build on the general literature on randomization inference for causal effects, (*e.g.* Neyman, 1990; Rosenbaum, 2002; Imbens and Rubin, 2015; Abadie et al., 2020; Rambachan and Roth, 2020). In particular, the discussion on the choice of estimands and its implications for randomization inference in Sekhon and Shem-Tov (2020) is relevant.

2 A MOTIVATING EXAMPLE

Synthetic Control estimators (Abadie et al., 2010) estimate treatment effects by comparing the outcome for the treated unit(s) to a weighted average of outcomes for control units, with weights chosen so that the weighted controls match the treated units prior to the treated period. The method requires data on outcomes for one or more treated units and multiple control units over a number of periods.

To preview the result that the SC estimator is biased, and that the bias can be removed



(a) Pre-treatment outcomes **(b)** Weights, SC estimator **(c)** Weights, USC estimator

Figure 1: Pre-treatment outcome in three-unit, two-period example. An outgoing arrow represents the weight assigned to that unit when the target of the arrow is treated, and arrows are colored by the unit the respective weight is put on.

by restricting the SC weights, we consider a simple three-unit (say, Arizona, AZ; California, CA; New York, NY), two-period ($t \in \{1, 2\}$) example where treatment is assigned to a single unit at time $t = 2$ with equal probability for each unit. To visually support this example, the three pre-treatment outcomes are depicted in Panel (a) of Figure 1. In this setting we compare the synthetic control (SC) estimator and an unbiased version of the SC estimator (USC).

Suppose CA is treated. Because CA's pre-treatment outcome is equidistant to each of AZ's and NY's outcome, the SC estimator puts equal weight on each of these control units. When either of the peripheral units, AZ or NY, is treated, the SC estimator puts all its weight on CA as putting any non-zero weights on the other unit would worsen the pre-treatment fit. The weights of the standard SC estimator are represented in Panel (b). We see from the graph that in expectation CA is used as a control unit more than AZ or NY, and, particularly, the total weight for CA as a control summed over all three assignments, which equals 2, exceeds the weight CA gets as a treated unit, which equals 1.

This difference in weights, or equivalently the imbalance of CA's use as treatment and control unit, is what creates bias in the SC estimator under randomization. Adding a simple constraint to the weights which rules out this imbalance yields an unbiased synthetic control estimator. The resulting weights can be found in Panel (c). In this three-unit example, the unbiased SC estimator is simply the difference-in-means estimator, but with more than three units this is not generally true.

3 SETUP

We consider a setting with N units, for which we observe outcomes Y_{it} for T time periods, $i = 1, \ldots, N$, $t = 1, \ldots, T$. There is a binary treatment that varies by units and time periods, denoted by $W_{it} \in \{0, 1\}$, and a pair of potential outcomes $Y_{it}(0)$ and $Y_{it}(1)$ for all unit/period combinations (Rubin, 1974; Imbens and Rubin, 2015). We assume there are no dynamic effects, so the potential outcomes are indexed only by the contemporaneous treatment. In some of the settings we consider, the dynamic effects would simply change the interpretation of the estimand. There are no restrictions on the time path of the potential outcomes. The $N \times T$ matrices of treatments and potential outcomes are denoted by $\boldsymbol{W}$, $\boldsymbol{Y}(0)$ and $\boldsymbol{Y}(1)$ respectively. Given the treatment, the realized/observed outcome matrix is $\boldsymbol{Y}$, with typical element $Y_{it} \equiv W_{it}Y_{it}(1) + (1 - W_{it})Y_{it}(0)$. In contrast to most of the SC literature (with Athey and Imbens, 2018 an exception), we take the potential outcomes $\boldsymbol{Y}(0)$ and $\boldsymbol{Y}(1)$ as fixed in our analysis, and treat the assignment matrix $\boldsymbol{W}$ (and thus the realized outcomes $\boldsymbol{Y}$) as stochastic.

For ease of exposition, we focus in much of the discussion on the case with a single treated unit and a single treated period. Many of the insights carry over to the case with a block of treated unit/time-period pairs, and we discuss explicitly the case with multiple treated units in Section 6.1. It is useful for our design-based analysis to separate out the assignment mechanism into the selection of the time period treated and the unit treated. For that purpose, exploiting the fact that there is only a single pair (i, t) with $W_{it} = 1$, we write

$$\boldsymbol{W} = \boldsymbol{U}\boldsymbol{V}^\top,$$

where $\boldsymbol{U}$ is an N-vector with typical element $U_i \in \{0, 1\}$ and $\sum_{i=1}^N U_i = 1$, and $\boldsymbol{V}$ is a T-vector with typical element $V_t \in \{0, 1\}$ and $\sum_{t=1}^T V_t = 1$, satisfying $V_t = \sum_{i=1}^N W_{it}$ and $U_i = \sum_{t=1}^T W_{it}$.

In many cases, the treated unit is exposed only in the last period, so $\boldsymbol{V}$ is non-stochastic, with last element equal to one and all other elements equal to zero. We examine this case separately, but additional insights are obtained by considering the more general case where both the treated unit and the treated period are stochastic.

3.1 Estimands

Next we define the estimands we consider in this article. Being precise about estimands is important for the discussion of bias and variance. It also clarifies the role of various assumptions we introduce.

Our primary focus is on the causal effect for the single treated unit/time-period:

$$\tau \equiv \tau(\boldsymbol{U}, \boldsymbol{V}) \equiv \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t \Big(Y_{it}(1) - Y_{it}(0) \Big). \tag{3.1}$$

For the case with multiple treated units or periods discussed in Section 6.1, this estimand can be generalized to the average effect for all the treated unit/time-periods. We focus on two properties of estimators for τ: the (exact, finite-sample) bias and variance.

There are three other estimands that are important to contrast with our primary estimand τ. First, the average effect for all N units in the treated period, which we call the "vertical" effect: $\tau^{\mathrm{V}} \equiv \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} V_t (Y_{it}(1) - Y_{it}(0))$. Second, the average effect for the treated unit over all T periods, which we call the "horizontal" effect: $\tau^{\mathrm{H}} \equiv \frac{1}{T} \sum_{i=1}^{N} \sum_{t=1}^{T} U_i (Y_{it}(1) - Y_{it}(0))$. Finally, the population average treatment effect: $\tau^{\mathrm{POP}} \equiv \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it}(1) - Y_{it}(0))$. If the unit and time period treated are both selected completely at random, then τ itself is unbiased for τ^{POP}, and by extension any unbiased estimator for τ is also unbiased for τ^{POP}. However, such an estimator has a larger variance as an estimator for τ^{POP} than as an estimator for τ.

Another advantage of focusing on τ rather than any of the other three estimands is that it frees us from conceptualizing $Y_{it}(1)$ for unit/time periods other than the unit/time pair that was actually treated. This can be important because in some cases it is difficult to give meaning to $Y_{it}(1)$ for units other than the treated unit. For example, in the German reunification application in Abadie et al. (2015), it is difficult to conceptualize $Y_{it}(1)$ for countries other than West Germany. Focusing on τ implies we only need to consider the value of West German GDP in the absence of the reunification.

3.2 Assumptions

In order to derive properties of the estimators, most of the SC literature uses a latent-factor model for the control outcome

$$Y_{it}(0) = \gamma_i' \beta_t + \varepsilon_{it} = \sum_{r=1}^{R} \gamma_{ir}\beta_{tr} + \varepsilon_{it},$$

here with R latent factors in combination with independence assumptions on the noise components ε_{it} (Abadie et al., 2010; Athey et al., 2017; Amjad et al., 2018; Xu, 2017). Here we focus instead on design assumptions, that is, assumptions about the assignment process that governs the distribution of $\boldsymbol{W}$ (or, equivalently, the distributions of $\boldsymbol{U}$ and $\boldsymbol{V}$). In this approach, we place no restrictions on the potential outcomes. Design-based, as opposed to model-based, approaches have been used in the general experimental design and program evaluation literature (*e.g.* Fisher, 1937; Neyman, 1990; Imbens and Rubin, 2015; Rosenbaum, 2002; Cunningham, 2018), as well as more recently in regression settings (Abadie et al., 2020; Athey and Imbens, 2018; Rambachan and Roth, 2020). However, to the best of our knowledge, these methods have not yet been used to analyze the properties of SC estimators.

First, we consider random assignment of the units to the treatment.

Assumption 1. (Random Assignment of Units)

$$\mathbb{P}(\boldsymbol{U} = \boldsymbol{u}) = \begin{cases} \frac{1}{N} & \text{if} \quad u_i \in \{0,1\}\ \forall i, \quad \sum_{i=1}^{N} u_i = 1, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{or} \quad \boldsymbol{U} \perp\!\!\!\perp (\boldsymbol{Y}(0), \boldsymbol{Y}(1)).$$

Although SC methods are typically used in observational studies, this assumption is the basis of the placebo tests that are often used in SC applications, *e.g.*, Abadie et al. (2010).

Second, we consider the assumption that the treated period was randomly selected from the T periods under observation.

Assumption 2. (Random Assignment of Treated Period)

$$\mathbb{P}(\boldsymbol{V} = \boldsymbol{v}) = \begin{cases} \frac{1}{T} & \text{if} \quad v_t \in \{0,1\}\ \forall t, \quad \sum_{t=1}^{T} v_t = 1, \\ 0 & \text{otherwise.} \end{cases} \qquad \text{or} \quad \boldsymbol{V} \perp\!\!\!\perp (\boldsymbol{Y}(0), \boldsymbol{Y}(1)).$$

Although this assumption is not plausible in many cases, as it is often only the last period(s) that are treated, it is useful to consider its implications. It formalizes the often implicit SC assumption that there is a within-period relationship between control outcomes for different units that is stable over time.

Most of our discussion concerns the finite-sample case. However, for some results it will be useful to consider large-T approximations. There are also other settings where it may be useful to consider large-N results, but we do not do so here. The analysis of large-N settings requires regularization of the SC weights because there are more weights to be estimated than periods to estimate them on, and the properties of the estimators will depend directly on the specific regularization method used (*e.g.* Abadie and L'Hour, 2017; Doudchenko and Imbens, 2016). For large-T results, first define $\boldsymbol{Y}_{\cdot t}(0)$ to be the N vector with typical element $Y_{it}(0)$. Define the averages up to period t of the first and the centered second moment:

$$\hat{\mu}_t = \frac{1}{t}\sum_{s=1}^{t} \boldsymbol{Y}_{\cdot s}(0), \qquad \hat{\Sigma}_t = \frac{1}{t}\sum_{s=1}^{t} \left(\boldsymbol{Y}_{\cdot s}(0) - \hat{\mu}_t\right)\left(\boldsymbol{Y}_{\cdot s}(0) - \hat{\mu}_t\right)^\top .$$

Assumption 3. (Large-T Stationarity) *For some finite μ and Σ the sequence of populations indexed by T satisfies, as $T \to \infty$,*

$$\hat{\mu}_T \longrightarrow \mu, \qquad \hat{\Sigma}_T \longrightarrow \Sigma.$$

3.3 Generalized Synthetic Control Estimators

In this section we introduce a class of SC type estimators. This class, which we refer to as Generalized Synthetic Control (GSC) estimators, includes the DiM estimator, the original SC estimator proposed by Abadie et al. (2010), as well as three modifications thereof as special cases. For the purpose of a randomization-based analysis, it is important that we define these estimators for all possible treatment assignment vectors $\boldsymbol{U}$ and $\boldsymbol{V}$, not just the realized assignment.

3.3.1 Estimators

We characterize the GSC estimators in terms of a set of weights $\boldsymbol{M}_{ijt}$, indexed by $i = 1, \ldots, N$, $j = 0, \ldots, N$, and $t = 1, \ldots, T$. Given a set of weights $\boldsymbol{M}$, treatment assignments

$\boldsymbol{U}$, $\boldsymbol{V}$, and outcomes $\boldsymbol{Y}$, the GSC estimator has the form

$$\hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}; \boldsymbol{M}) \equiv \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t \left\{ \boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt} \right\}. \tag{3.2}$$

If unit i is treated in period t, the estimator is a linear function of the outcomes for all units for period t. For all estimators we consider, the weights are a function of the baseline outcomes $\boldsymbol{Y}(0)$ during non-treated periods.

The estimators in the GSC class differ in the choice of the weights $\boldsymbol{M}$. There are generally two components to this choice. First, there is a set of possible weights, denoted by $\mathcal{M}$, over which we search for an optimal weight. These sets $\mathcal{M}$ differ between the estimators we consider, but in all cases the sets are non-stochastic. Second, there is an objective function that defines the weight within the set of possible weights. This objective function is identical for all GSC estimators we consider in the current article.

We only consider weights in the following set:

$$\mathcal{M}^0 = \left\{ \boldsymbol{M} \,\middle|\, \boldsymbol{M}_{iit} = 1, \forall i \geq 1, t; \boldsymbol{M}_{ijt} \leq 0, \forall i \geq 1, t; \sum_{j=1}^{N} \boldsymbol{M}_{ijt} = 0 \, \forall i \geq 1, t \right\}. \tag{3.3}$$

There are three restrictions captured in this set. First, the weight for the treated unit is equal to one. Second, the weight for unit j for the prediction of the causal effect for unit i is nonpositive:

$$\boldsymbol{M}_{ijt} \leq 0, \;\; \forall i \in \{1, \ldots, N\}, j \in \{1, \ldots, N\} \setminus \{i\}, t \in \{1, \ldots, T\}. \tag{3.4}$$

The third restriction requires that the weights for all units in the prediction for the causal effect for unit i in period t sum to zero. Because the weight for unit i in this prediction is restricted to be equal to one, this means that the weights for the control units sum to minus one.

We consider four estimators in this class, characterized by four sets of possible weights $\mathcal{M} \subset \mathcal{M}_0$, described in Section 3.3.3.

3.3.2 The Objective Function

We start with the second component of the choice of weights, the objective function. For a given matrix of outcomes $\boldsymbol{Y}$, and a given set of possible weights $\mathcal{M}$, define the tensor $\boldsymbol{M}(\boldsymbol{Y};\mathcal{M})$ with elements $\boldsymbol{M}_{ijt}$, as:

$$\boldsymbol{M}(\boldsymbol{Y};\mathcal{M}) \equiv \arg\min_{\boldsymbol{M}\in\mathcal{M}} \sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\sum_{s\neq t}\left(\boldsymbol{M}_{i0t}+\sum_{j=1}^{N}\boldsymbol{M}_{ijt}Y_{js}\right)^2\right\}. \tag{3.5}$$

This leads to the following characterization of the GSC estimator as a function of $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}$ and the set of weights $\mathcal{M}$:

$$\hat{\tau}^{\text{GSC}} = \sum_{i=1}^{N}\sum_{t=1}^{T} U_i V_t \left\{\boldsymbol{M}_{i0t}+\sum_{j=1}^{N}\boldsymbol{M}_{ijt}Y_{jt}\right\}, \tag{3.6}$$

where

$$\boldsymbol{M} = \arg\min_{\boldsymbol{M}\in\mathcal{M}} \sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\sum_{s\neq t}\left(\boldsymbol{M}_{i0t}+\sum_{j=1}^{N}\boldsymbol{M}_{ijt}Y_{js}\right)^2\right\}. \tag{3.7}$$

What is the motivation for the objective function in (3.5)? The expected squared error of the estimator $\hat{\tau}^{\text{GSC}}$, under unit and time randomization (Assumptions 1 and 2), is

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\boldsymbol{M}_{i0t}+\sum_{j=1}^{N}\boldsymbol{M}_{ijt}Y_{jt}(0)\right)^2.$$

We cannot evaluate this squared loss because it depends on $Y_{it}(0)$ that we do not observe. However, we can use the analogue from the control periods, both before and after the treated period. Under time randomization, these values are comparable to the current value, suggesting the objective function (3.5).

3.3.3 Feasible Weights

The DiM estimator corresponds to the case with

$$\mathcal{M}^{\text{DiM}} = \left\{\boldsymbol{M}\in\mathcal{M}^0 \middle| \boldsymbol{M}_{i0t}=0\,\forall i,t, M_{ijt}=-1/(N-1)\,\forall i\neq j,t\right\}.$$

Relative to DiM, the DiD estimator relaxes the no-intercept restriction,

$$\mathcal{M}^{\text{DiD}} = \left\{ \boldsymbol{M} \in \mathcal{M}^0 \middle| M_{ijt} = -1/(N-1) \, \forall i \neq j, t \right\}.$$

The original SC estimator (Abadie and Gardeazabal, 2003; Abadie et al., 2010) corresponds to the estimator in (A.1) with the set $\mathcal{M}$ defined as the subset of $\mathcal{M}^0$ satisfying

$$\mathcal{M}^{\text{SC}} = \left\{ \boldsymbol{M} \in \mathcal{M}^0 \middle| \boldsymbol{M}_{i0t} = 0 \, \forall i, t \right\}.$$

The modification introduced in Doudchenko and Imbens (2016) and Ferman and Pinto (2019), allows for an intercept by dropping the restriction $\boldsymbol{M}_{i0t} = 0$, leading to the Modified Synthetic Control (MSC) estimator with $\mathcal{M}^{\text{MSC}} = \mathcal{M}^0$. Arkhangelsky et al. (2019) show that the inclusion of the intercept can be interpreted as including a unit fixed effect in the regression function. In Section 4 we discuss how the inclusion of the intercept ties in with the time randomization assumption. The presence of the intercept also reduces the importance of time-invariant covariates.

A second modification of the basic SC estimator, the Unbiased Synthetic Control (USC) estimator, places an additional set of restrictions on the weights, namely that all units are in expectation used as controls as often as they are used as treated units:

$$\mathcal{M}^{\text{USC}} = \left\{ \boldsymbol{M} \in \mathcal{M}^0 \middle| \boldsymbol{M}_{i0t} = 0, \, \forall i, t, \sum_{i=1}^{N} \boldsymbol{M}_{ijt} = 0 \, \forall j \geq 1, t \right\}.$$

Finally, we combine the two modifications of the SC estimator, the inclusion of the intercept and the additional restriction, leading to our main alternative to the SC estimator, the Modified Unbiased Synthetic Control (MUSC) estimator:

$$\mathcal{M}^{\text{MUSC}} = \left\{ \boldsymbol{M} \in \mathcal{M}^0 \middle| \sum_{i=1}^{N} \boldsymbol{M}_{ijt} = 0 \, \forall j \geq 1, t \right\}. \tag{3.8}$$

These four sets of restrictions define four estimators. Our focus is primarily on $\hat{\tau}^{\text{SC}}$ and $\hat{\tau}^{\text{MUSC}}$, with the comparison with $\hat{\tau}^{\text{MSC}}$ and $\hat{\tau}^{\text{USC}}$ serving primarily to aid the interpretation of the two restrictions that make up the difference between $\hat{\tau}^{\text{SC}}$ and $\hat{\tau}^{\text{MUSC}}$.

4 PROPERTIES

In this section, we investigate the properties of the various estimators given the unit and/or time randomization assumptions.

4.1 The Bias of the SC Estimator

The first question we study is the bias of the four SC estimators relative to the treatment effect for the treated unit, τ. We summarize the results in Table 2.

Table 2: Bias Properties of SC Estimators for τ (B is biased, U is unbiased)

	Maintained Assumptions				
Estimator $\downarrow$	Unit Randomization	Time Randomization	Time Randomization Large T	Unit & Time Randomization	Unit & Time Randomization Large T
$\hat{\tau}^{\text{SC}}$	B	B	B	B	B
$\hat{\tau}^{\text{MSC}}$	B	B	U	B	U
$\hat{\tau}^{\text{USC}}$	U	B	B	U	U
$\hat{\tau}^{\text{MUSC}}$	U	B	U	U	U

The basic SC estimator is not unbiased even if both the unit treated and the time period treated were both randomly selected. Because this is perhaps surprising, it is useful to study the bias of the SC estimator in more detail. Recall the general definition of the SC type estimators in (3.2). We can rewrite this as

$$\hat{\tau}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{Y},\boldsymbol{M}) = \sum_{i=1}^{N}\sum_{t=1}^{T} U_i V_t \left\{ \boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt} \right\}$$

$$= \sum_{t=1}^{T}\sum_{i=1}^{N} V_t U_i \left\{ Y_{it}(1) + \boldsymbol{M}_{i0t} + \sum_{j=1}^{N} (1-U_j)\boldsymbol{M}_{ijt} Y_{jt}(0) \right\}.$$

We focus on the properties relative to our primary estimand, the treatment effect for the treated. The estimation error is equal to

$$\hat{\tau}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{Y},\boldsymbol{M}) - \tau(\boldsymbol{U},\boldsymbol{V}) = \sum_{i=1}^{N}\sum_{t=1}^{T} U_i V_t \left\{ Y_{it}(0) + \boldsymbol{M}_{i0t} + \sum_{j=1}^{N} (1-U_j)\boldsymbol{M}_{ijt} Y_{jt}(0) \right\}.$$

We consider the bias of the SC estimators separately for the estimators without an intercept (the SC and USC estimators), and for the estimators with the intercept estimated through minimization of the objective function (3.5) (the MSC and MUSC estimators).

Lemma 1. *Suppose Assumption 1 (random assignment of units to treatment) holds. Then if one of the following two conditions holds:*
(i) the intercept is zero, $\boldsymbol{M}_{i0t} = 0$ for all i, t,
or (ii) if the intercept is estimated through (3.5),
then the conditional (on $\boldsymbol{V}$) bias vanishes if the set of weights $\mathcal{M}$ guarantees that

$$\sum_{i=1}^{N} \boldsymbol{M}_{ijt} = 0 \;\; \forall j = 1, \ldots, N, t = 1, \ldots, T. \tag{4.1}$$

This lemma shows that τ^{USC} and τ^{MUSC} are unbiased, because both estimators only search over weight sets that satisfy the adding-up condition in (4.1).

For the SC estimator, the conditional bias under Assumption 1 is

$$\text{Bias}^{\text{SC}} = \mathbb{E}\left[\hat{\tau} - \tau \middle| \boldsymbol{V}\right] = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} V_t Y_{it}(0) \sum_{j=1}^{N} \boldsymbol{M}_{jit}^{\text{SC}}.$$

The intuition for the bias of the SC estimator also holds for the simple matching estimator (*e.g.*, Abadie and Imbens, 2006), which is generally biased under randomization in finite samples.

In principle one can estimate the bias for the SC estimator and generate an unbiased estimator by subtracting the estimated bias from the standard SC estimator. However, the properties of this de-biased estimator are not very attractive in terms of RMSE.

To see the role that time randomization plays in the bias, consider the MSC estimator in a setting with large T, and random selection of the treated period. For ease of exposition, suppose unit N is the treated unit. As noted before, one can view the MSC estimator as a regression estimator where we regress the outcomes $Y_{N1}, \ldots, Y_{NT}$ on the treatment indicator and the predictors $Y_{1,t}, \ldots, Y_{N-1,t}$ and an intercept. It is well known that this leads to an estimator that is asymptotically unbiased in large samples (in this case meaning large T; Freedman, 2008; Lin, 2013; Imbens and Rubin, 2015). This does not affect the bias of the SC estimator because it does not have an intercept.

A final comment concerns the magnitude of the bias. In the illustration in the Introduction, and in some of the simulations below, the bias of the SC estimator is modest. This raises the question whether there are any guarantees that the bias of the SC estimator is small. This is not the case, and the bias can in fact be arbitrarily large. Consider a case with binary outcomes, $Y_{it} \in \{0,1\}$, with the treatment effect equal to zero for all units. Suppose the number of time periods is equal to the number of units. Moreover, suppose that for the first unit $Y_{1t} = 0$ for $t = 1, \ldots, T$. For all units $i \neq 1$ $Y_{it} = 0$ for $t \notin \{i, T\}$ with $Y_{ii} = Y_{iT} = 1$. In that case the first unit is matched with equal weight to all other units, $M_{ijT} = -1/(N-1)$ for j=2, $\ldots, N-1$, and all other units are matched to the first unit: $M_{i1T} = -1$ for $j = 2, \ldots, N$. The bias in this case is $(N-2)/N$ which can be made arbitrarily close to the maximum possible value of 1.

4.2 A Network Interpretation of SC Estimators and Their Bias

To understand the bias of SC estimators, we note that SC weights

$$\boldsymbol{M} \in \mathcal{M}^{\mathrm{SC}} = \left\{ \boldsymbol{M} \,\middle|\, \boldsymbol{M}_{iit}{=}1, \forall i, t; \boldsymbol{M}_{ijt}{\leq}0, \forall i, j{\geq}1, t; \sum_{j=1}^{N} \boldsymbol{M}_{ijt} = 0 \; \forall i, t; \boldsymbol{M}_{i0t}{=}0 \forall i \right\}$$

for a given treatment time t can be understood as a directed network with vertices i and edge flows (or weights) $w_{ij} = -\boldsymbol{M}_{jit} \geq 0$ from vertex i to vertex $j \neq i$. The weight constraint $\sum_{j=1}^{N} \boldsymbol{M}_{ijt} = 0$ then ensures that the total incoming flow equals one for all vertices i, $\sum_{j \neq i} w_{ij} = 1$. An example of such a network representation of an SC estimator is given in Figure 1 (b).

Bias can arise in the SC network whenever the incoming flow (which measures how often a unit is treated) is not the same as the outgoing flow of a vertex (which measures how often each unit is used as a control). The network corresponding to the SC estimator in Figure 1 is imbalanced: for the outside vertices, inflow exceeds outflow, while the inside vertex has higher outflow than inflow. Imposing the unbiasedness constraint $\sum_{j=1}^{N} \boldsymbol{M}_{jit} = 0$ is equivalent to imposing the flow balance constraint

$$\sum_{j \neq i} w_{ij} = \sum_{j \neq i} w_{ji}$$

at all vertices i, ensuring that the corresponding units are used as often as controls as they are treated. Such a network is obtained in Figure 1 (c), where inflows and outflows are balanced for all vertices.

Beyond providing an intuitive language to represent synthetic-control estimators, we show in Section 6.3 how tools from network analysis can help analyzing their properties. There, we show that the eigenvector centrality in the network represented by W relates to propensity scores subject to which an SC estimator is unbiased. With this network representation, we thus connect the SC estimator to the tools and insights from the literature on networks across statistics and the social sciences (*e.g.*, Jackson, 2010; de Paula, 2017, 2020).

4.3 The Exact GSC Variance and its Unbiased Estimation

Consider the GSC estimator $\hat{\tau} = \hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M})$ as an estimator of τ.

Lemma 2. *Suppose Assumption 1 holds. Then*

$$\mathbb{V}(\boldsymbol{V}, \boldsymbol{M}) = \mathbb{E}\left[(\hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}) - \tau)^2 \middle| \boldsymbol{V} \right] = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} V_t \left(\boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt}(0) \right)^2 .$$

For the unbiased estimators, this is the variance, while for the other estimators, it is the expected squared error. The challenge is that the variance depends on control outcomes that we do not observe. However, we can estimate this variance without bias under unit-randomization.

Proposition 1. *Suppose Assumption 1 holds. Then the estimator*

$$\begin{aligned}\hat{\mathbb{V}} = \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t \Bigg\{ & \frac{1}{N-3} \sum_{\substack{k=1 \\ k \neq i}}^{N} \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{M}_{kjt} (Y_{kt} - Y_{jt}) \right)^2 - \frac{1}{(N-2)(N-3)} \sum_{\substack{k=1 \\ k \neq i}}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{M}_{kjt}^2 (Y_{kt} - Y_{jt})^2 \\ & + \frac{2}{N-2} \sum_{\substack{k=1 \\ k \neq i}}^{N} \boldsymbol{M}_{k0t} \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{M}_{kjt} (Y_{jt} - Y_{kt}) \right) + \frac{1}{N} \sum_{k=1}^{N} \boldsymbol{M}_{k0t}^2 \Bigg\} ,\end{aligned} \tag{4.2}$$

is unbiased for $\mathbb{V}(\boldsymbol{V}, \boldsymbol{M})$.

This result may be somewhat surprising. Note that in completely randomized experiments there is no unbiased estimator of the variance of the simple difference-in-means estimator for the average treatment effect (Imbens and Rubin, 2015). The current result is different because here we focus on the average effect for the treated only. For that case, there is an unbiased estimator for the variance of the difference-in-means estimator in the case of randomized experiments (*e.g.*, Sekhon and Shem-Tov, 2020).

The variance estimator in this proposition has three terms. The first takes the form of a leave-one-out estimator based on the control units excluding the treated unit. The remaining two terms correct for over-counting the diagonal elements in the inner square of the first term and additional terms for the intercept. In the special case of the DiM estimator the variance reduces to the standard one and in that case it is guaranteed to be nonnegative.

4.4 Improvement over the Difference-in-Means Estimator

In Section 5.2 we show that the variance of some of the GSC estimators can be substantially smaller than that of the DiM estimator. However, that is not guaranteed under the assumption that the treated unit was randomly selected. It is possible that in the treated period the pattern between the outcomes is very different from that in the other periods, so that the GSC estimators have a variance that is larger than that of the DiM estimator. However, this can be ruled out if the treated time period is randomly selected among all periods (Assumption 2). Formally, we need the number of time periods to be large for this to hold. Suppose unit N is treated. Define

$$\hat{\beta}_N = \arg\min_{\beta} \sum_{t=1}^{T} (Y_{Nt} - \beta_0 - \sum_{j=1}^{N-1} \beta_j Y_{jt})^2.$$

Now let β_N^* be the limit of β_N as $T \to \infty$. Then the expected squared error of using the GSC estimator based on β_N^* for a randomly selected time period is less than using the DiM estimator:

$$\mathbb{E}\left[\sum_{t=1}^{T} V_t \left(Y_{Nt}(0) - \beta_0^* - \sum_{j=1}^{N-1} \beta_j^* Y_{jt}\right)^2\right] \leq \mathbb{E}\left[\sum_{t=1}^{T} V_t \left(Y_{Nt}(0) - \sum_{j=1}^{N-1} Y_{jt}/(N-1)\right)^2\right].$$

Moreover, with T large, this also holds for

$$\hat{\beta}_N = \arg\min_{\beta} \sum_{t=1}^{T} (1 - V_t)(Y_{Nt} - \beta_0 - \sum_{j=1}^{N-2} \beta_j Y_{jt})^2,$$

under standard assumptions that guarantee that $\hat{\beta}_N \to \beta_N^*$. Finally, this is true for any unit i, implying that with large T the variance for the GSC estimator will not be larger than for the DiM estimator.

4.5 The Placebo Variance Estimator

To put the proposed variance estimator in (4.2) in perspective, we consider here an alternative approach for estimating the variance of the SC estimator. Versions of this placebo variance estimator have been proposed previously both for testing zero effects (*e.g.* Abadie et al., 2010) and for constructing confidence intervals (*e.g.* Doudchenko and Imbens, 2016). Suppose unit i is the treated unit. We put this unit aside, and focus on the $N-1$ control units. For each of these $N-1$ control units (indexed by $j = 1, \ldots, N, j \neq i$), we recalculate the weights, leaving out the treated unit, and then estimate the treatment effect. For ease of exposition we focus on the case where the last period is the treated period, $V_T = 1$.

We now define $(N-1) \times N$ weight matrices and sets of weight matrices $\boldsymbol{M}^{(i)}$ and $\mathcal{M}^{(i)}$. For $\boldsymbol{M}^{(i)}$ we have $\boldsymbol{M}^{(i)}_{ij} = \boldsymbol{M}^{(i)}_{ji} = 0$ for all j. The weights are defined as

$$\boldsymbol{M}^{(i)}(\boldsymbol{Y}, \mathcal{M}^{(i)}) = \underset{\boldsymbol{M}^{(i)} \in \mathcal{M}^{(i)}}{\arg\min} \sum_{s=1}^{T-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(\boldsymbol{M}^{(i)}_{j0} + \sum_{\substack{k=1 \\ k \neq i}}^{N} \boldsymbol{M}^{(i)}_{jk} Y_{ks} \right)^2 . \tag{4.3}$$

Given these weights, the placebo estimator is

$$\hat{\tau}^{(i)}_j = \boldsymbol{M}^{(i)}_{j0} + \sum_{\substack{k=1 \\ k \neq i}}^{N} \boldsymbol{M}^{(i)}_{jk} Y_{kT}, \qquad \text{for } j \neq i.$$

Because unit j is a control unit, this is an estimator of zero, and the placebo variance estimator uses it to estimate the variance of $\hat{\tau}$ as

$$\hat{\mathbb{V}}^{PCB} = \frac{1}{N-1} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} U_i \left(\hat{\tau}^{(i)}_j \right)^2 = \frac{1}{N-1} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} U_i \left(\boldsymbol{M}^{(i)}_{j0} + \sum_{\substack{k=1 \\ k \neq i}}^{N} \boldsymbol{M}^{(i)}_{jk} Y_{kT} \right)^2 .$$

The point we wish to make here is that this variance estimator can be upward as well as downward biased, depending on the potential outcomes. In order to demonstrate this, we give two examples in the Appendix, one where the placebo variance estimator is biased downward, and one where it is biased upward.

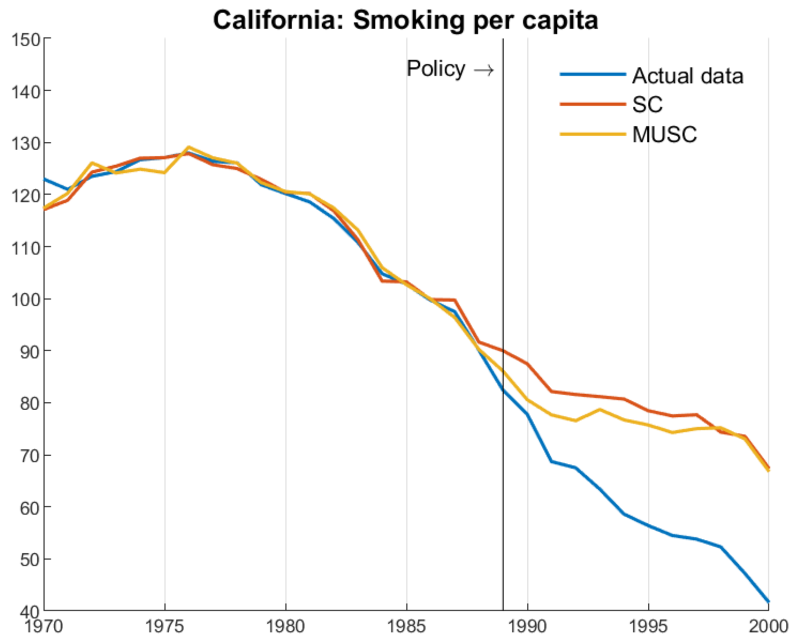


Figure 2: Pre- and Post-Treatment fit of SC and MUSC

5 AN ILLUSTRATION AND SOME SIMULATIONS

In this section we illustrate some of the methods proposed in this article.

5.1 The California Smoking Study

To illustrate some of the concepts developed in this article, we first turn to the data from the California smoking study (Abadie et al., 2010). In Figure 2 we compare the SC and MUSC estimates. We find that the pre-treatment fit is similar for both estimators despite the additional restriction. In addition, the point estimates are similar. The interpretation is that the number of "similar" control units is large enough that a single additional restriction (and the relaxing of another one) does not affect the goodness of fit substantially.

5.2 A Simulation Study

We also perform a small simulation study to assess the properties of the MUSC estimator. Following Bertrand et al. (2004) and Arkhangelsky et al. (2019), we use data from the Current Population Survey for $N = 50$ states and $T = 40$ years. The variables we analyze include state/year average log wages, hours, and the state/year unemployment rate. The true treatment effects are all zero by construction. This allows us to calculate the RMSE. For each of the variables, we estimate the treatment effects using the Difference-in-Means (DiM) estimator, the standard Synthetic Control (SC) estimator, the Difference-in-Differences (DiD) estimator, and the Modified Unbiased Synthetic Control (MUSC) estimator. We also include for comparison the LASSO estimator based on regressing the period T outcomes on all the lagged outcomes, using ℓ_1 regularization. The LASSO estimator is not an SC-type estimator and is generally biased in our framework, and we merely include it as a benchmark that uses more information than the DiM estimator. For the LASSO estimator, we choose the regularization parameter using leave-one-out cross validation.

The first simulation study we conduct compares the performance of the four estimators in terms of RMSE. The study is designed as follows. For each treated period $T \geq 21$, we use $T - 1$ pre-treatment periods to estimate the weights for the different estimators and discard all data after T. Then, iterating through all 50 states, we pretend that each state has been selected for treatment and calculate the corresponding estimated treatment effects. Lastly, we average over all states to summarize the performance in terms of RMSE for a single treated period. In Table 3, we report the results for averaged over all twenty years. We report the results for the setting with 50 units, as well as for settings with 10 and 5 units to assess the relative performance with fewer cross-sectional units. We find that the RMSE is substantially lower for the SC and MUSC estimator compared to the DiM estimator for all variables. In Table 5 in the Appendix, we document that these results hold across years. The SC and MUSC estimators perform comparably to the LASSO estimator for log wages and outperform it for the other two variables in the case with $N = 50$, with the DiD estimator performing worse than any of these three. Note that the SC and MUSC estimators have similar RMSE in this case. In the cases with $N = 10$ and $N = 5$, the MUSC estimator substantially outperforms the other methods except for the DiD estimator, which performs slightly worse for $N = 10$ and slightly better for $N = 5$. Overall, the MUSC estimator performs consistently well, while the relative performance of LASSO and DiD vary substantially with the number of units.

The second simulation study demonstrates the properties of our proposed unbiased variance estimator and the placebo variance estimator. Here we focus on the average log wages

Table 3: Simulation Experiment Based on CPS Data Averaged over States and Years – RMSE

Est	$N = 50$			$N = 10$			$N = 5$		
	Ln(Wage)	Hours	U-Rate	Ln(Wage)	Hours	U-Rate	Ln(Wage)	Hours	U-Rate
DiM	0.105	1.197	0.015	0.122	1.558	0.016	0.144	1.380	0.015
SC	0.051	0.918	0.013	0.070	1.045	0.015	0.079	1.010	0.014
MUSC	0.053	0.903	0.013	0.061	0.960	0.015	0.057	0.944	0.013
LASSO	0.051	0.952	0.013	0.077	1.313	0.017	0.116	1.251	0.015
DiD	0.063	0.976	0.013	0.066	0.982	0.014	0.055	0.929	0.013

DiM: Difference-in-Means estimator, SC: Synthetic Control estimator of Abadie et al. (2010), MUSC: Modified Unbiased Synthetic Control estimator, DiD: Difference-in-Differences estimator.

and fix $T = 40$ as the treated period. Moreover, we decrease the sample to $N = 20$ units in total. For ease of readability, we choose to report the standard error based on our variance estimator instead of the variances themselves. Table 4 reports standard errors based on the true variance along with the estimates (averaged over all units) based on our variance estimator and using the placebo approach. We find that our estimator is indeed unbiased. The placebo approach is very modestly biased – the direction of the bias depends on which estimator is used. For the DiM, the DiD, and the SC estimator, the placebo estimator is upward biased; for the MUSC estimator it is downward biased.

Table 4: Simulation Experiment Based on CPS Data (Log Wages) by State and Year for $N = 20$ units and Treated Period $T = 40$ – Average Standard Error

	$\sqrt{V_{\text{true}}}$	$\sqrt{\hat{V}_{\text{GSC}}}$	$\sqrt{\hat{V}_{\text{PCB}}}$
DiM	0.1157	0.1157	0.1158
SC	0.0518	0.0518	0.0521
MUSC	0.0490	0.0490	0.0489
DiD	0.0553	0.0553	0.0554

DiM: Difference in Means, SC: Synthetic Control, MUSC: Modified Synthetic Control, DiD: Difference in Differences, GSC: General Synthetic Control, PCB: Placebo

6 GENERALIZATIONS AND EXTENSIONS

In this section, we present three generalizations of the setting considered so far. The focus up to now was on the case with a single treated unit and single treated period where the estimand was the average effect for the treated under random assignment. First, we consider the case with multiple treated units. Second, we consider the case where the estimand is the average effect for all units in the treated period. Both of these generalizations create conceptual complications. Third, we consider the case of non-constant propensity scores.

6.1 Multiple Treated Units

In this section, we look at the case with multiple treated units. We fix the number of treated units at N_{T}. The estimand is, as before, the average effect for the N_{T} treated units:

$$\tau = \tau(\boldsymbol{U}, \boldsymbol{V}) \equiv \frac{1}{N_{\mathrm{T}}} \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t \Big(Y_{it}(1) - Y_{it}(0) \Big).$$

We modify Assumption 1 to

Assumption 4. (Random Assignment of Units)

$$\mathbb{P}(\boldsymbol{U} = \boldsymbol{u}) = \begin{cases} \left(\frac{N!}{N_{\mathrm{T}}! N_{\mathrm{C}}!} \right)^{-1} & \text{if} \quad u_i \in \{0,1\} \, \forall i, \quad \sum_{i=1}^{N} u_i = N_{\mathrm{T}}, \\ 0 & \text{otherwise.} \end{cases}$$

In this case, it is convenient to work with the sets of units assigned to the treatment, rather than the individual units. There are $K = N!/(N_{\mathrm{T}}!N_{\mathrm{C}}!) = \binom{N}{N_T}$ such sets. Of these, $(N-1)!/((N_{\mathrm{T}}-1)!N_{\mathrm{C}}!) = \binom{N-1}{N_T-1}$ include a given unit, such as unit 1, since there are $\binom{N-1}{N_T-1}$ combinations of the remaining units if that unit is treated. This represents a fraction N_{T}/N of the total number of sets of N_{T} treated units.

Let $\tilde{\boldsymbol{U}}$ be the vector of length K of indicators denoting which set of N_{T} units is treated. Let e_k be the K-component vector with the k-th component equal to one and all other elements equal to zero. By construction, $\sum_{k=1}^{K} \tilde{\boldsymbol{U}}_k = 1$, and $\tilde{\boldsymbol{U}}_k \in \{0,1\}$. Assumption 4 implies that the probability that $\tilde{\boldsymbol{U}}_k = 1$ is equal to $1/K$. Let $u_i(\tilde{\boldsymbol{U}}) \in \{0,1\}$ be an indicator for unit i being treated given the assignment vector $\tilde{\boldsymbol{U}}$. In this notation, we can rewrite τ as

$$\tau = \tau(\tilde{\boldsymbol{U}}, \boldsymbol{V}) = \frac{1}{N_{\mathrm{T}}} \sum_{k=1}^{K} \sum_{t=1}^{T} \tilde{\boldsymbol{U}}_k V_t \sum_{i=1}^{N} u_i(\tilde{\boldsymbol{U}}_k) \Big(Y_{it}(1) - Y_{it}(0) \Big).$$

Instead of the tensors $\boldsymbol{M}$ with dimension $N \times (N+1) \times T$, we now have tensors with dimension $K \times (N+1) \times T$, with one row for each of the $K = N!/(N_{\mathrm{T}}!N_{\mathrm{C}}!)$ possible sets of treated units. The estimators we consider are of the form

$$\hat{\tau}(\tilde{\boldsymbol{U}}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}) \equiv \sum_{k=1}^{K} \sum_{t=1}^{T} \tilde{\boldsymbol{U}}_k V_t \left\{ \boldsymbol{M}_{k0t} + \sum_{j=1}^{N} \boldsymbol{M}_{kjt} Y_{jt} \right\}. \tag{6.1}$$

This formulation suggests the restriction that $\boldsymbol{M}_{kjt} = 1/N_T$, for all j such that $u_j(e_k) = 1$ and $\boldsymbol{M}_{kjt} \leq 0$, whenever $u_j(e_k) = 0$. The set of such $\boldsymbol{M}$ we consider for the generalized modified unbiased synthetic control (MUSC) estimator is

$$\mathcal{M}^{\text{MUSC}} = \left\{ \boldsymbol{M} \middle| \sum_{j=1}^{N} \boldsymbol{M}_{kjt} = 0 \, \forall k, t, \sum_{k=1}^{K} \boldsymbol{M}_{kjt} = 0 \, \forall j \geq 1, t \right\}.$$

The objective function for choosing $\boldsymbol{M}$ is now

$$\boldsymbol{M}(\boldsymbol{Y}, \mathcal{M}^{\text{MUSC}}) = \underset{\boldsymbol{M} \in \mathcal{M}^{\text{MUSC}}}{\arg\min} \sum_{k=1}^{K} \sum_{t=1}^{T} \left\{ \sum_{s \neq t} \left(\boldsymbol{M}_{k0t} + \sum_{j=1}^{N} \boldsymbol{M}_{kjt} Y_{js} \right)^2 \right\}.$$

Lemma 3. *Suppose that Assumption 4 holds. Then*
(i) the estimator $\hat{\tau}(\tilde{\boldsymbol{U}}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M}^{\text{MUSC}}))$ *is unbiased conditional on* $\boldsymbol{V}$*:*

$$\mathbb{E}\left[\hat{\tau}(\tilde{\boldsymbol{U}}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M}^{\text{MUSC}})) - \tau(\boldsymbol{U}, \boldsymbol{V}) \middle| \boldsymbol{V} \right] = 0,$$

(ii) the variance of $\hat{\tau}(\tilde{\boldsymbol{U}}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M}^{\text{MUSC}})$ *is*

$$\mathbb{V}\left(\hat{\tau}(\tilde{\boldsymbol{U}}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M}^{\text{MUSC}})) \middle| \boldsymbol{V} \right) = \frac{1}{K} \sum_{t=1}^{T} V_t \sum_{k=1}^{K} \left(\boldsymbol{M}_{k0t} + \sum_{j=1}^{N} \boldsymbol{M}_{kjt} Y_{ts}(0) \right)^2,$$

and (iii), the variance can be estimated without bias (conditional on $\boldsymbol{V}$*) by a generalization of the variance estimator in Proposition 1,*

$$\begin{aligned}
\hat{\mathbb{V}} = \sum_{k=1}^{K} \sum_{t=1}^{T} \tilde{\boldsymbol{U}}_k V_t \Bigg(\sum_{\substack{k'=1; \\ u_i(\tilde{\boldsymbol{U}}_k) + u_i(\tilde{\boldsymbol{U}}_{k'}) \leq 1 \forall i}}^{K} \Bigg\{ & \frac{1}{\binom{N_C - 2}{N_T}} \left(\sum_{j=1}^{N} (1 - u_j(\tilde{\boldsymbol{U}}_k)) \boldsymbol{M}_{k'jt} (Y_{jt} - \overline{Y}_{k't}) \right)^2 \\
& - \frac{N_T}{(N_C - 1)\binom{N_C - 2}{N_T}} \sum_{j=1}^{N} (1 - u_j(\tilde{\boldsymbol{U}}_k)) \boldsymbol{M}_{k'jt}^2 (Y_{jt} - \overline{Y}_{k't})^2 \\
& + \frac{2}{\binom{N_C - 1}{N_T}} \boldsymbol{M}_{k'0t} \sum_{j=1}^{N} (1 - u_j(\tilde{\boldsymbol{U}}_k)) \boldsymbol{M}_{k'jt} (Y_j - \overline{Y}_{k'}) \Bigg\} + \frac{1}{K} \sum_{k'=1}^{K} \boldsymbol{M}_{k'0t}^2 \Bigg)
\end{aligned}$$

for $\overline{Y}_{k't} = \frac{1}{N_T} \sum_{j=1}^{N} u_j(\tilde{\boldsymbol{U}}_{k'}) Y_{jt}$.

6.2 The Average Effect for All Units

Here we look at the case where the estimand changes from the average effect for the treated unit(s) to the average effect over all units in the treated periods. For ease of exposition, we continue to focus on the case with a single treated period and a single treated unit. Formally, the estimand is

$$\tau^{\mathrm{V}} = \tau^{\mathrm{V}}(\boldsymbol{V}) \equiv \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} V_t\Big(Y_{it}(1) - Y_{it}(0)\Big),$$

We can separate this into two components, the effect for the treated unit, and the average effect for the controls

$$\tau^{\mathrm{T}} \equiv \sum_{i=1}^{N}\sum_{t=1}^{T} U_i V_t\Big(Y_{it}(1) - Y_{it}(0)\Big), \quad \tau^{\mathrm{C}} \equiv \frac{1}{N-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(1-U_i)V_t\Big(Y_{it}(1) - Y_{it}(0)\Big),$$

with $\tau^{\mathrm{V}} = \tau^{\mathrm{T}}/N + \tau^{\mathrm{C}}(N-1)/N$. Consider, as before, an estimator of the form

$$\hat{\tau}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{Y},\boldsymbol{M}) = \sum_{i=1}^{N}\sum_{t=1}^{T} U_i V_t \left\{ \boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt} \right\}.$$

The restrictions $\boldsymbol{M}_{iit} = 1\ \forall i,t$, $\sum_{i=1}^{N} \boldsymbol{M}_{ijt} = 0\ \forall j,t$ (including the intercept) still imply unbiasedness conditional on $\boldsymbol{V}$, and the MUSC remains unbiased for τ^{V}. Yet the variance (and more generally the conditional expected loss of such a weighted estimator) is now

$$\mathbb{E}\left[\left(\hat{\tau}(\boldsymbol{U},\boldsymbol{V},\boldsymbol{Y},\boldsymbol{M}) - \tau^{V}\right)^2 \Big| \boldsymbol{V}\right] = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} V_t\Bigg(\boldsymbol{M}_{i0t} + \left(\boldsymbol{M}_{iit} - \frac{1}{N}\right) Y_{it}(1)$$
$$- \sum_{j=1}^{N}(1-U_j)\frac{1}{N}Y_{jt}(1) + \frac{1}{N}Y_{it}(0) + \sum_{j=1}^{N}(1-U_j)\left(\boldsymbol{M}_{ijt} + \frac{1}{N}\right) Y_{jt}(0)\Bigg)^2 \tag{6.2}$$

which depends on treated and untreated potential outcomes. This creates two related challenges. First, since the expression depends on treated outcomes, there is no immediate sample analogue available that corresponds to minimizing expected error, even under time randomization. Second, the variance cannot generally be estimated without bias, since it depends not only on the variation of the $Y_{it}(0)$ (which can be estimated), but also on the variation of the $Y_{it}(1)$ and their covariance with the $Y_{it}(0)$ (neither of which is identified from the data).

We briefly discuss two possibilities for the (non-stochastic) correlation of treatment and control outcomes, and what they imply for estimation. First, if treatment effects are constant within time period (and treatment and control potential outcomes thus perfectly correlated, $Y_{it}(1) - Y_{it}(0) = \tau$), then (6.2) becomes

$$\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} V_t \left(\boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt}(0) \right)^2$$

as before, suggesting the MUSC estimator. If, on the other hand, treated outcomes are uncorrelated to control outcomes, then (6.2) becomes

$$\frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} V_t \left(\boldsymbol{M}_{i0t} + \frac{1}{N} \sum_{j=1}^{N} Y_{jt}(0) + \sum_{j=1}^{N} (1 - U_j) \boldsymbol{M}_{ijt} Y_{jt}(0) \right)^2 + \text{const.},$$

which suggests an alternative MUSC-type estimator that minimizes the sample analogue in non-treated time periods over weights $\mathcal{M}^{\text{MUSC}}$, which could effectively shrink the MUSC weights on control units towards the DiM weights.

6.3 Non-Constant Propensity Scores

So far we have assumed that treatment is assigned with equal probability across units, time periods, or unit–time pairs. Yet the theory we develop generalizes to non-constant propensity scores. Here, we focus on the case where treatment happens at time t and is assigned randomly to single unit i with probability $p_i^{(t)} \in [0, 1]$, where $\sum_{i=1}^{N} p_i^{(t)} = 1$. We ask whether an estimator $\hat{\tau}^{\text{GSC}}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}; \boldsymbol{M}) \equiv \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t \left\{ \boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt} \right\}$ is unbiased for $\tau(\boldsymbol{U}, \boldsymbol{V}) = \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t Y_{it}(1) - Y_{it}(0)$ with respect to these propensity scores.

Proposition 2. *The estimator $\hat{\tau}^{\text{GSC}}$ is unbiased for $\tau(\boldsymbol{U}, \boldsymbol{V})$ (across values of potential outcomes) if and only if $\boldsymbol{M} \in \mathcal{M}_p^{\text{MUSC}} = \left\{ \boldsymbol{M} \middle| \sum_{j=1}^{N} \boldsymbol{M}_{ijt} = 0 \, \forall i, t, \sum_{i=1}^{N} p_i^{(t)} \boldsymbol{M}_{ijt} = 0 \, \forall j, t \right\}$.*

Here, unbiasedness generalizes the adding-up condition $\sum_{i=1}^{N} \boldsymbol{M}_{ijt} = 0$ from the class of MUSC matrices to its propensity-weighted analogue $\sum_{i=1}^{N} p_i^{(t)} \boldsymbol{M}_{ijt} = 0$, which ensures that the bias is zero since $\mathbb{E}\left[\hat{\tau} - \tau \middle| \boldsymbol{V}\right] = \sum_{t=1}^{T} V_t \left(\sum_{i=1}^{N} Y_{it}(0) \sum_{j=1}^{N} p_j^{(t)} \boldsymbol{M}_{jit} + \sum_{j=1}^{N} p_j^{(t)} \boldsymbol{M}_{j0t} \right)$. A natural analogue of the MUSC estimator is then

$$\boldsymbol{M}_p^{\text{MUSC}}(\boldsymbol{Y}, \mathcal{M}_p^{\text{MUSC}}) \equiv \underset{\boldsymbol{M} \in \mathcal{M}_p^{\text{MUSC}}}{\arg\min} \sum_{i=1}^{N} p_i^{(t)} \sum_{t=1}^{T} \left\{ \sum_{s \neq t} \left(\boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{js} \right)^2 \right\}. \qquad (6.3)$$

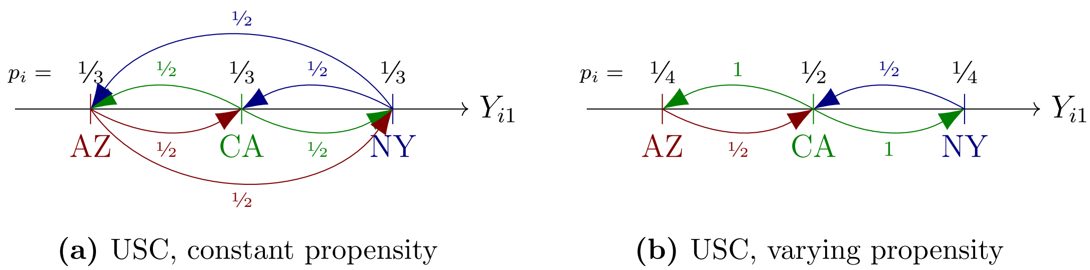


Figure 3: Pre-treatment outcome in three-unit, two-period example with varying treatment propensities. An outgoing arrow represents the weight assigned to that unit when the target of the arrow is treated, and arrows are colored by the unit the respective weight is put on.

Such an estimator could be used when treatment is assigned randomly. Note also that the variance estimator from Proposition 1 extends. When the analyst has a choice over the treatment assignment, and $t = T$, the optimization in (6.3) could also include the choice of propensity score.

We now illustrate how varying propensity scores can affect the synthetic control estimator, extending the motivating example in Section 2. In this example, we consider the standard SC estimator, the USC estimator with equal propensities, and the USC estimator with non-constant propensities. Recall that the standard SC estimator is biased under this design and the USC estimator corrects this bias by enforcing balance between the probability of being treated and being used as a control. However, when the central unit is treated with higher probability of ½ (Figure 3b), then the weight matrix that only uses the closest units as control in each case is the optimal unbiased solution. In this specific example, this solution also coincides with the standard SC solution.

In this example, there is a set of propensity scores for which the standard SC estimator is unbiased. This is not a coincidence. As the following proposition shows, for the weights of every SC-type estimator there is a set of propensity scores such that the corresponding estimator is unbiased.

Proposition 3. *For every SC weight matrix*

$$\boldsymbol{M} \in \mathcal{M}^{\mathrm{SC}} = \left\{ \boldsymbol{M} \,\middle|\, \boldsymbol{M}_{iit}{=}1, \forall i, t; \boldsymbol{M}_{ijt}{\leq}0, \forall i, j{\geq}1, t; \sum_{j=1}^{N} \boldsymbol{M}_{ijt} = 0 \; \forall i, t; \boldsymbol{M}_{i0t}{=}0 \forall i \right\}$$

there exists a propensity-score vector $p^{(t)} \in [0,1]^n, \sum_{i=1}^{N} p_i = 1$ such that the corresponding *estimator is unbiased with respect to $p^{(t)}$ at treatment time t.*

This proposition does not rely on the weight matrix $\boldsymbol{M}$ being the result of a specific optimization program, as it applies to any weight matrix that follows the basic structure of the SC matrices (without intercept).

To understand the propensity scores $p^{(t)}$ that make the SC estimator associated with the weights $\boldsymbol{M}$ unbiased, we note that they can be interpreted as a measure of centrality in the network associated with $\boldsymbol{M}$ in a precise way, where more central units correspond to higher propensity scores. Specifically, let $W = (w_{ij})_{i,j \in \{1,\ldots,N\}}$ be the edge flow matrix from Section 4.2 corresponding to $\boldsymbol{M}$ for treatment time t, meaning that $w_{ij} = -\boldsymbol{M}_{jit}$ for $i \neq j$ and $w_{ii} = 0$. Then the eigenvector centralities of vertices in this network are equivalent (up to normalization) to the propensity scores that ensure unbiasedness, where we consider the case where both are unique.

Proposition 4. *Assume that the network W associated with $\boldsymbol{M}$ for treatment at t is strongly connected (equivalently, that W is irreducible). Then the propensity score vector $p^{(t)}$ for which the estimator is unbiased at t is unique and the same as the eigenvector centrality in the network W (with appropriate normalization).*

This connection between eigenvectors and unbiased propensities follows naturally from the representation of the estimator in terms of its (weighted) network adjacency matrix W. Writing $M = (\boldsymbol{M}_{ijt})_{i,j \in \{1,\ldots,N\}}$ for the $N \times N$ matrix corresponding to the SC weights when treatment happens at t, the unbiasedness condition corresponds to $M' p^{(t)} = \mathbf{0}$. Since $M' = \mathbb{I} - W$, $p^{(t)}$ is an eigenvector of W with eigenvalue 1, which is also the largest eigenvalue and corresponds to the unique non-negative eigenvector if the network is strongly connected. When the network is not strongly connected, we may still obtain a similar result for its components.

These results suggest ways in which considering varying propensity scores can be helpful when analyzing SC-type estimators. First, when treatment is randomized according to a known probability distribution, then those probabilities affect the optimal USC and MUSC weights. Second, the propensities that make an estimator unbiased have an intuitive interpretation as the eigenvector centralities of the network corresponding to the weight matrix of a synthetic-control estimator. Third, even in the observational case, varying propensities could be used when some units can be considered to be more likely to receive treatment or to be more appropriate as controls, allowing to replace binary inclusion criteria by treatment propensities. Finally, when we choose propensity scores in the design of an experiment and plan to use an SC-type estimator, then we can optimize the choice of propensities based on past outcomes to be better suited to their relationship, assigning more central units higher probabilities.

7 CONCLUSION

In this article we study Synthetic Control (SC) methods from a design perspective. We show that when a randomized experiment is conducted, the standard SC estimator is biased. However, a minor modification of the SC estimator is unbiased under randomization, and in cases with few treated units can have RMSE properties superior to those of the standard Difference-in-Means estimator. We show that the design perspective also has implications for observational studies. We propose a variance estimator that is validated by randomization.

References

Abadie, A. (2019). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*.

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2020). Sampling-based versus design-based uncertainty in regression analysis. *Econometrica*, 88(1):265–296.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, pages 495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(-):113–132.

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Abadie, A. and L'Hour, J. (2017). A penalized synthetic control estimator for disaggregated data. *Work. Pap., Mass. Inst. Technol., Cambridge, MA.*

Amjad, M., Shah, D., and Shen, D. (2018). Robust synthetic control. *The Journal of Machine Learning Research*, 19(1):802–852.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2019). Synthetic difference in differences. Technical report, National Bureau of Economic Research.

Athey, S., Bayati, M., Doudchenko, N., Imbens, G., and Khosravi, K. (2017). Matrix completion methods for causal panel data models. *arXiv preprint arXiv:1710.10251.*

Athey, S. and Imbens, G. W. (2018). Design-based analysis in difference-in-differences settings with staggered adoption. Technical report, National Bureau of Economic Research.

Ben-Michael, E., Feller, A., and Rothstein, J. (2020). The augmented synthetic control method. *arXiv preprint arXiv:1811.04170.*

Berman, A. and Plemmons, R. J. (1994). 2. Nonnegative Matrices. In *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, pages 26–62. Society for Industrial and Applied Mathematics.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Cavallo, E., Galiani, S., Noy, I., and Pantano, J. (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*, 95(5):1549–1561.

Chernozhukov, V., Wuthrich, K., and Zhu, Y. (2017). An exact and robust conformal inference method for counterfactual and synthetic controls. *arXiv preprint arXiv:1712.09089.*

Coffman, M. and Noy, I. (2012). Hurricane iniki: measuring the long-term economic impact of a natural disaster using synthetic control. *Environment and Development Economics*, 17(2):187–205.

Cunningham, S. (2018). *Causal inference: The mixtape.* Yale University Press.

de Paula, Á. (2017). Econometrics of Network Models. In *Advances in Economics and Econometrics: Eleventh World Congress*, pages 268–323. Cambridge University Press.

de Paula, Á. (2020). Econometric Models of Network Formation. *Annual review of economics*, 12(1):775–799.

Doudchenko, N. and Imbens, G. W. (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. Technical report, National Bureau of Economic Research.

Ferman, B. and Pinto, C. (2017). Placebo tests for synthetic controls.

Ferman, B. and Pinto, C. (2019). Synthetic controls with imperfect pre-treatment fit. *arXiv preprint arXiv:1911.08521.*

Firpo, S. and Possebom, V. (2018). Synthetic control method: Inference, sensitivity analysis and confidence sets. *Journal of Causal Inference*, 6(2).

Fisher, R. A. (1937). *The design of experiments.* Oliver And Boyd; Edinburgh; London.

Freedman, D. A. (2008). On regression adjustments in experiments with several treatments. *The annals of applied statistics*, 2(1):176–196.

Hahn, J. and Shi, R. (2016). Synthetic control and inference. *Available at UCLA.*

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences.* Cambridge University Press.

Jackson, M. O. (2010). *Social and Economic Networks.* Princeton University Press.

Lei, L. and Candès, E. J. (2020). Conformal inference of counterfactuals and individual treatment effects. *arXiv preprint arXiv:2006.06138.*

Li, K. T. (2020). Statistical inference for average treatment effects estimated by synthetic control methods. *Journal of the American Statistical Association*, 115(532):2068–2083.

Lin, W. (2013). Agnostic notes on regression adjustments for experimental data: Reexamining freedman's critique. *The Annals of Applied Statistics*, 7(1):295–318.

Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, 87:25–41.

Neyman, J. (1923/1990). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 5(4):465–472.

Rambachan, A. and Roth, J. (2020). Design-Based Uncertainty for Quasi-Experiments. *arXiv preprint arXiv:2008.00602*.

Rosenbaum, P. R. (2002). *Observational Studies*. Springer.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.

Sekhon, J. S. and Shem-Tov, Y. (2020). Inference on a new class of sample average treatment effects. *Journal of the American Statistical Association*, pages 1–7.

Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76.

Table 5: Simulation Experiment Based on CPS Data by State and Year – RMSE

Treated	Log Wages					Hours					Unemployment Rate				
Period	DiM	SC	MUSC	LASSO	DiD	DiM	SC	MUSC	LASSO	DiD	DiM	SC	MUSC	LASSO	DiD
$T = 21$	0.1157	0.0543	0.0550	0.0560	0.0636	1.4838	0.9797	0.9340	1.1999	1.1063	0.0123	0.0111	0.0102	0.0110	0.0114
$T = 22$	0.1112	0.0496	0.0502	0.0512	0.0708	1.3764	0.8204	0.9170	0.9963	1.0276	0.0107	0.0086	0.0084	0.0102	0.0096
$T = 23$	0.1089	0.0435	0.0484	0.0409	0.0540	1.2821	0.9091	0.9009	0.9904	0.9197	0.0114	0.0111	0.0116	0.0107	0.0110
$T = 24$	0.1143	0.0421	0.0450	0.0452	0.0626	1.2400	0.9569	0.8829	0.9524	0.9747	0.0145	0.0146	0.0142	0.0145	0.0148
$T = 25$	0.1136	0.0465	0.0500	0.0446	0.0602	1.3125	0.9074	0.8896	0.9011	1.0567	0.0124	0.0110	0.0115	0.0125	0.0114
$T = 26$	0.1111	0.0487	0.0471	0.0509	0.0639	1.1424	0.8509	0.8347	0.8456	0.8544	0.0132	0.0136	0.0146	0.0132	0.0123
$T = 27$	0.1105	0.0481	0.0493	0.0439	0.0506	1.1926	0.8612	0.8141	0.9746	0.9063	0.0155	0.0145	0.0149	0.0155	0.0139
$T = 28$	0.1014	0.0551	0.0674	0.0597	0.0640	1.0627	0.7611	0.7660	0.8451	0.9322	0.0136	0.0105	0.0106	0.0122	0.0099
$T = 29$	0.0964	0.0473	0.0594	0.0487	0.0628	1.1342	0.8552	0.8686	0.8797	1.0101	0.0137	0.0124	0.0118	0.0133	0.0109
$T = 30$	0.0980	0.0505	0.0516	0.0562	0.0626	1.0741	0.7765	0.7755	0.7318	0.8350	0.0137	0.0141	0.0147	0.0137	0.0140
$T = 31$	0.1084	0.0410	0.0393	0.0429	0.0691	1.2340	0.9700	0.9537	0.9405	1.0748	0.0209	0.0180	0.0175	0.0170	0.0196
$T = 32$	0.1049	0.0460	0.0468	0.0475	0.0600	1.0571	0.8496	0.8209	0.8495	0.8910	0.0214	0.0182	0.0167	0.0180	0.0177
$T = 33$	0.1046	0.0530	0.0536	0.0485	0.0594	1.2947	1.0018	0.9696	1.0183	1.0211	0.0215	0.0161	0.0151	0.0133	0.0175
$T = 34$	0.1060	0.0625	0.0662	0.0610	0.0741	1.1640	0.9174	0.9007	0.9836	1.0024	0.0182	0.0131	0.0129	0.0122	0.0142
$T = 35$	0.1053	0.0560	0.0537	0.0529	0.0623	1.2233	1.0824	0.9914	1.0680	0.9688	0.0201	0.0137	0.0137	0.0166	0.0179
$T = 36$	0.0985	0.0594	0.0535	0.0530	0.0593	1.1267	0.8083	0.7396	0.8650	0.9133	0.0150	0.0124	0.0119	0.0134	0.0121
$T = 37$	0.1017	0.0605	0.0579	0.0548	0.0672	1.3066	1.2080	1.1235	1.2517	1.0831	0.0156	0.0128	0.0131	0.0145	0.0126
$T = 38$	0.0929	0.0580	0.0615	0.0607	0.0665	0.9917	0.7742	0.7929	0.7677	0.8234	0.0126	0.0116	0.0117	0.0113	0.0108
$T = 39$	0.0853	0.0459	0.0554	0.0435	0.0628	0.8437	0.7979	0.9475	0.6663	0.9233	0.0112	0.0120	0.0122	0.0112	0.0114
$T = 40$	0.1051	0.0517	0.0479	0.0527	0.0598	1.4048	1.2714	1.2382	1.3024	1.1896	0.0126	0.0112	0.0106	0.0114	0.0103
Average	0.1047	0.0510	0.0530	0.0507	0.0628	1.1974	0.9180	0.9031	0.9515	0.9757	0.0150	0.0130	0.0129	0.0132	0.0132

DiM: Difference-in-Means estimator, SC: Synthetic Control estimator of Abadie et al. (2010), MUSC: Modified Unbiased Synthetic Control estimator, DiD: Difference-in-Differences estimator.

A ADDITIONAL RESULTS

A.1 Placebo Variance Examples

Here we present examples for which the placebo variance is biased. For simplicity, we consider the case where treatment occurs in the last of $T = 3$ periods and there are $N = 4$ units. In this etting, the placebo variance can be biased downward:

Example 1. *Suppose that for some arbitrary* a, b, c, d*,*

$$\boldsymbol{Y}(0) = \begin{pmatrix} a & b & 0 \\ a & b & 1 \\ c & d & 0 \\ c & d & 1 \end{pmatrix}, \quad \text{leading to} \quad \boldsymbol{M}^{\text{MUSC}} = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix},$$

so that the units are matched in pairs. If unit 1 *is treated, the estimator is* $Y_{13} - Y_{23}$*, with error* $Y_{13}(0) - Y_{23}(0) = -1$*. Similar calculations for the other three units show that the squared error is always equal to 1, and hence the true variance is 1.*

Now let us calculate the placebo variance. Here we exploit the fact that with three units the weights for all units are equal. This leads to

$$\boldsymbol{M}^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ -((a+b)-(c+d))/2 & 0 & 1 & -1/2 & -1/2 \\ -((c+d)-(a+b))/4 & 0 & -1/2 & 1 & -1/2 \\ -((c+d)-(a+b))/4 & 0 & -1/2 & -1/2 & 1 \end{pmatrix},$$

Then the placebo variance is smaller in expectation than the true variance.

In the same setting, the placebo variance can be biased upward:

Example 2. *Suppose that*

$$\boldsymbol{M} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \qquad \boldsymbol{Y}_{\cdot T} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

so the units are matched in pairs, and the matching is of perfect quality. Then the placebo variance is higher in expectation than the true variance.

A.2 Non-Stochastic Weights

Here, we argue formally that we can consider the weights to be non-stochastic for the main analysis in our article. For a given set $\mathcal{M}$ we can write each estimator as

$$\hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M})) = \sum_{i=1}^{N} \sum_{t=1}^{T} U_i V_t \left\{ \boldsymbol{M}_{i0t}(\boldsymbol{Y}, \mathcal{M}) + \sum_{j=1}^{N} \boldsymbol{M}_{ijt}(\boldsymbol{Y}, \mathcal{M}) Y_{jt} \right\}. \tag{A.1}$$

For this class of estimators we can view the weights as non-stochastic:

Lemma A.1. *For all* $\boldsymbol{Y}(0)$, $\boldsymbol{Y}(1)$, $\boldsymbol{U}$, $\boldsymbol{V}$, *and* $\mathcal{M}$,

$$\hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M})) = \hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}(0), \mathcal{M})) \tag{A.2}$$

This representation is useful because the properties of $\hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}(0), \mathcal{M}))$ are easier to establish under assumptions on $\boldsymbol{V}$ and $\boldsymbol{U}$ than those of $\hat{\tau}(\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{Y}, \mathcal{M}))$ for the general case.

B PROOFS

Proof of Proposition 1. We first consider the case without intercept. As a preliminary calculation, note that for $k, j, j' \in \{1, \ldots, n\}$

$$\sum_{i=1}^{N} \sum_{k,j,j' \neq i} \frac{1}{N - |\{k, j, j'\}|} a_{kjj'} = \sum_{k,j,j'} a_{kjj'}, \tag{A.1}$$

since every term kjj' term appears $N - |\{k, j, j'\}|$ times in the sum on the left. Let now

$$a_{kjj'} = M_{kj}(Y_j(0) - Y_k(0)) \cdot M_{kj'}(Y_{j'}(0) - Y_k(0)),$$

where for simplicity we fix the period t, drop all time indices to set $M_{ij} = \boldsymbol{M}_{ijt}$, and write $\hat{\mathbb{V}}_i$ for the variance estimator when $U_i = 1$. Then $a_{kjj'} = 0$ for $k \in \{j, j'\}$ and thus

$$\frac{1}{N} \sum_{i=1}^{N} \underbrace{\left(\frac{1}{N-3} \sum_{k \neq i} \left(\sum_{j \neq i} M_{kj}(Y_j(0) - Y_k(0)) \right)^2 - \frac{1}{(N-3)(N-2)} \sum_{k,j \neq i} M_{kj}^2 (Y_j(0) - Y_k(0))^2 \right)}_{=\hat{\mathbb{V}}_i}$$

$$= \frac{1}{N} \sum_{i=1}^{N} \Big(\sum_{k,j,j' \neq i} \frac{1}{N-3} a_{kjj'} - \sum_{\substack{k,j,j' \neq i \\ j = j'}} \underbrace{\frac{1}{(N-3)(N-2)}}_{= \frac{1}{N-3} - \frac{1}{N-2}} a_{kjj'} \Big)$$

$$= \frac{1}{N} \Big(\sum_{i=1}^{N} \sum_{\substack{k,j,j' \neq i \\ |\{k,j,j'\}| = 3}} \frac{1}{N-3} a_{kjj'} + \sum_{\substack{k,j,j' \neq i \\ |\{k,j,j'\}| = 2}} \frac{1}{N-2} \underbrace{a_{kjj'}}_{=0 \text{ for } j \neq j'} + \sum_{\substack{k,j,j' \neq i \\ |\{k,j,j'\}| = 1}} \frac{1}{N-1} \underbrace{a_{kjj'}}_{=0} \Big)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \sum_{k,j,j' \neq i} \frac{1}{N - |\{k, j, j'\}|} a_{kjj'} \overset{\text{(A.1)}}{=} \frac{1}{N} \sum_{k,j,j'} a_{kjj'}$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} M_{ij}(Y_j(0) - Y_i(0)) \right)^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} M_{ij} Y_j(0) \right)^2 = \mathbb{V}.$$

Here, we have used that $\sum_{j=1}^{N} M_{ij} = 0$.

With an intercept we note that

$$\mathbb{V} = \frac{1}{N}\sum_{i=1}^{N}\left(M_{i0} + \sum_{j=1}^{N} M_{ij}Y_j(0)\right)^2 = \frac{1}{N}\sum_{i=1}^{N}\left(M_{i0} + \sum_{j=1}^{N} M_{ij}(Y_j(0) - Y_i(0))\right)^2$$
$$= \frac{1}{N}\sum_{i=1}^{N} M_{i0}^2 + \frac{2}{N}\sum_{i=1}^{N} M_{i0}\left(\sum_{j=1}^{N} M_{ij}(Y_j(0) - Y_i(0))\right) + \frac{1}{N}\left(\sum_{j=1}^{N} M_{ij}(Y_j(0) - Y_i(0))\right)^2,$$

where

$$\frac{2}{N-2}\sum_{k\neq i} M_{k0}\left(\sum_{j\neq i} M_{kj}(Y_j(0) - Y_k(0))\right)$$

is unbiased for the middle term, using that $M_{kj}(Y_j(0) - Y_k(0)) = 0$ for $k = j$. It follows that

$$\hat{\mathbb{V}}_i = \frac{1}{N-3}\sum_{k\neq i}\left(\sum_{j\neq i} M_{kj}(Y_j(0) - Y_k(0))\right)^2 - \frac{1}{(N-3)(N-2)}\sum_{k,j\neq i} M_{kj}^2(Y_j(0) - Y_k(0))^2$$
$$+ \frac{2}{N-2}\sum_{k\neq i} M_{k0}\left(\sum_{j\neq i} M_{kj}(Y_j(0) - Y_k(0))\right) + \frac{1}{N}\sum_{k} M_{k0}^2$$

is an unbiased estimator of the conditional variance $\mathbb{V}$. □

Proof of the variance expression in Lemma 3. This proof generalized the proof of Proposition 1 above. Specifically, for $[N] = \{1, \ldots, N\}$,

$$\sum_{\substack{k\subseteq[N];\\|k|=N_T}}\ \sum_{\substack{i\subseteq[N]\setminus k;\\|i|=N_T}}\ \sum_{j,j'\in[N]\setminus k\cup\{0\}} \frac{1}{\binom{|[N]\setminus(k\cup[\{j,j'\})|}{N_T}} a_{i,j,j'} = \sum_{\substack{k\subseteq\{1,\ldots,N\};\\|k|=N_T}}\ \sum_{j,j'\in[N]\cup\{0\}} a_{k,j,j'} \tag{A.2}$$

for a conformal tensor a.

For fixed t as above consider weights M_{kj} indexed by $k \subseteq [N]$ with $|k| = N_T$ and $j \in [N] \cup 0$, for which (a) $\sum_{j=1}^{N} M_{kj} = 0$ and (b) $M_{kj} = 1$ for $j \in k$, and potential outcomes $Y_j(0)$ with $j \in [N]$. Write $\overline{Y}_k(0) = \frac{1}{N_T}\sum_{j\in k} Y_j(0)$. (This approach generalizes to cases where treated units are themselves weighted, in which case we would replace $\overline{Y}_k(0)$ by the corresponding weighted average.) Let

$$b_{k,j} = \begin{cases} 0, & j \in k, \\ M_{kj}(Y_j(0) - \overline{Y}_k(0)), & j \in [N] \setminus k, \\ M_{k0}, & j = 0, \end{cases} \qquad a_{k,j,j'} = b_{k,j}b_{k,j'}. \tag{A.3}$$

Then, for $K = \binom{N}{N_T}$, and using that (c) $a_{k,j,j'} = 0$ whenever j or j' are in $[N] \setminus k$ and (d) $a_{k,j,j'} = a_{k,j',j}$,

$$
\begin{aligned}
\mathbb{V} &= \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \left(M_{k0} + \sum_{j=1}^{N} M_{kj} Y_j(0) \right)^2 \overset{(a)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \left(M_{k0} + \sum_{j=1}^{N} M_{kj} (Y_j(0) - \overline{Y}_k(0)) \right)^2 \\
&\overset{(b)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \left(M_{k0} + \sum_{j \in [N] \setminus k} M_{kj} (Y_j(0) - \overline{Y}_k(0)) \right)^2 \overset{(A.31)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \left(\sum_{j=0}^{N} b_{kj} \right)^2 \\
&\overset{(A.32)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \sum_{j,j' \in [N] \cup \{0\}} a_{k,j,j'} \overset{(A.2)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \sum_{\substack{i \subseteq [N] \setminus k; \\ |i| = N_T}} \sum_{j,j' \in [N] \setminus k \cup \{0\}} \frac{1}{\binom{|[N] \setminus (k \cup [\{j,j'\}])|}{N_T}} a_{i,j,j'} \\
&\overset{(c)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \sum_{\substack{i \subseteq [N] \setminus k; \\ |i| = N_T}} \left(\frac{a_{i,0,0}}{\binom{N_C}{N_T}} + \sum_{j \in [N] \setminus k} \frac{a_{i,j,0} + a_{i,0,j} + a_{i,j,j}}{\binom{N_C - 1}{N_T}} + \sum_{\substack{j,j' \in [N] \setminus k \\ j \neq j'}} \frac{a_{i,j,j'}}{\binom{N_C - 2}{N_T}} \right) \\
&\overset{(d)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \left(a_{k,0,0} + \sum_{\substack{i \subseteq [N] \setminus k; \\ |i| = N_T}} \left(\sum_{j \in [N] \setminus k} \left(\frac{2 a_{i,j,0}}{\binom{N_C - 1}{N_T}} - \underbrace{\frac{N_T}{(N_C - 1)\binom{N_C - 2}{N_T}}}_{= \frac{1}{\binom{N_C - 2}{N_T}} - \frac{1}{\binom{N_C - 1}{N_T}}} a_{i,j,j} \right) + \sum_{j,j' \in [N] \setminus k} \frac{a_{i,j,j'}}{\binom{N_C - 2}{N_T}} \right) \right) \\
&\overset{(A.32)}{=} \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \left(b_{k,0}^2 + \sum_{\substack{i \subseteq [N] \setminus k; \\ |i| = N_T}} \left(\frac{2 b_{i,0} \sum_{j \in [N] \setminus k} b_{i,j}}{\binom{N_C - 1}{N_T}} - \frac{N_T \sum_{j \in [N] \setminus k} b_{i,j}^2}{(N_C - 1)\binom{N_C - 2}{N_T}} + \frac{\left(\sum_{j \in [N] \setminus k} b_{i,j} \right)^2}{\binom{N_C - 2}{N_T}} \right) \right) \\
&\overset{(A.31)}{=} \frac{1}{K} \left(\frac{1}{K} \sum_{\substack{i \subseteq [N]; \\ |i| = N_T}} M_{i0}^2 + \sum_{\substack{i \subseteq [N] \setminus k; \\ |i| = N_T}} \left\{ \frac{2}{\binom{N_C - 1}{N_T}} M_{i0} \sum_{j \in [N] \setminus k} M_{ij} (Y_j - \overline{Y}_i) \right. \right. \\
&\qquad \left. \left. + \frac{1}{\binom{N_C - 2}{N_T}} \left(\sum_{j \in [N] \setminus k} M_{ij} (Y_j - \overline{Y}_i) \right)^2 - \frac{N_T}{(N_C - 1)\binom{N_C - 2}{N_T}} \sum_{j \in [N] \setminus k} M_{ij}^2 (Y_j - \overline{Y}_i)^2 \right\} \right) \\
&= \frac{1}{K} \sum_{\substack{k \subseteq [N]; \\ |k| = N_T}} \hat{\mathbb{V}}_k,
\end{aligned}
$$

so the proposed estimator is unbiased for the variance.

An alternative estimator that emphasizes the leave-fold-out nature of this construction is

$$\sum_{\substack{i \subseteq [N]\setminus k;\\ |i|=N_T}} \left\{ \frac{1}{\binom{N_C}{N_T}} M_{i0}^2 + \frac{2}{\binom{N_C-1}{N_T}} M_{i0} \sum_{j\in[N]\setminus k} M_{ij}(Y_j - \overline{Y}_i) \right.$$
$$\left. + \frac{1}{\binom{N_C-2}{N_T}} \left(\sum_{j\in[N]\setminus k} M_{ij}(Y_j - \overline{Y}_i) \right)^2 - \frac{N_T}{(N_C-1)\binom{N_C-2}{N_T}} \sum_{j\in[N]\setminus k} M_{ij}^2 (Y_j - \overline{Y}_i)^2 \right\}.$$

It has the additional advantage that it does not use the weights on the treated observations when constructing the variance for a specific draw. Here, the first average of squared intercepts could be replaced by the overall average of squared intercepts, as in the main variance estimator above. □

Proof of Proposition 2. The result immediately follows from

$$\begin{aligned} \mathbb{E}\left[\hat{\tau} - \tau \,|\, \boldsymbol{V}\right] &= \sum_{t=1}^{T} V_t \sum_{i=1}^{N} p_i^{(t)} \left(\boldsymbol{M}_{i0t} + \sum_{j=1}^{N} \boldsymbol{M}_{ijt} Y_{jt}(0) \right) \\ &= \sum_{t=1}^{T} V_t \left(\sum_{j=1}^{N} p_j^{(t)} \boldsymbol{M}_{j0t} + \sum_{i=1}^{N} Y_{it}(0) \sum_{j=1}^{N} p_j^{(t)} \boldsymbol{M}_{jit} \right). \end{aligned}$$ □

Proof of Proposition 3. The existence is a consequence of the existence of non-negative eigenvectors in the associated network (which can be established e.g. by the Perron–Frobenius theorem). For a direct proof, note that the weight matrix $M \in \mathbb{R}^{N\times N}$ defined by $M_{ij} = \boldsymbol{M}_{ijt}$, $i,j \in \{1,\ldots,N\}$, fulfills

$$M_{ij} \begin{cases} = 1, & i = j, \\ \leq 0, & i \neq j \end{cases}, \qquad M\mathbf{1} = \mathbf{0}. \tag{A.4}$$

By (A.4) M is singular, so there exists some vector $q \in \mathbb{R}^N \setminus \{\mathbf{0}\}$ such that $M'q = \mathbf{0}$. Define $p \in [0,1]^N, p'\mathbf{1} = 1$ by $p_i = |q_i| / \sum_j |q_j|$. If q has only non-negative or only non-positive values, we are done. Assume now that q has both positive and negative elements q_i, and denote by I^- the indices corresponding to negative elements, and by I^+ the indices corresponding to positive elements. Then

$$0 = \sum_{i\in I^-} \underbrace{(M'q)_i}_{=0} = \underbrace{\sum_{i\in I^-}\sum_{j\in I^-} M'_{ij} q_j}_{=\sum_{j\in I^-} \underbrace{q_j}_{<0} \underbrace{\sum_{i\in I^-} M'_{ij}}_{\geq 0} \leq 0} + \sum_{i\in I^-}\sum_{j\in I^+} \overbrace{M'_{ij}}^{\leq 0} \underbrace{q_j}_{>0} \leq 0$$

and thus $\sum_{i \in I^-} M'_{ij} = 0$ for all $j \in I^-$ and $M'_{ij} = 0$ for all $(i,j) \in I^- \times I^+$. Since $\sum_i M'_{ij} = 0$ and $M'_{ij} \leq 0$ whenever $i \neq j$, it follows from $\sum_{i \in I^-} M'_{ij} = 0$ that $M'_{ij} = 0$ for all $i \notin I^-$, $j \in I^-$. We therefore have that

$$\left(\sum_j |q_j|\right)(M'p)_i = \left(\sum_j |q_j|\right)\left(\sum_{j \in I^-} M'_{ij} p_j + \sum_{j \in I^+} M_{ij} p_j\right)$$
$$= \begin{cases} -\sum_{j \in I^-} M'_{ij} q_j, & i \in I^-, \\ \sum_{j \in I^+} M'_{ij} q_j, & i \notin I^- \end{cases} = \begin{cases} -(M'q)_i, & i \in I^-, \\ (M'q)_i, & i \notin I^- \end{cases} = 0$$

and thus $M'p = \mathbf{0}$. □

Proof of Proposition 4. By Perron–Frobenius (e.g. Berman and Plemmons, 1994), for W non-negative and irreducible, W has a unique (up to scaling) non-negative eigenvector (which is positive), and that eigenvector corresponds to the largest eigenvalue of W. Since p in Proposition 3 is non-negative and an eigenvector with eigenvalue 1, is is the same as the eigenvector centrality of the network W. □