On the Assumptions of Synthetic Control Methods

Claudia Shi
Columbia University

Dhanya Sridhar
University of Montreal

Vishal Misra
Columbia University

David M. Blei
Columbia University

Abstract

Synthetic control (SC) methods have been widely applied to estimate the causal effect of large-scale interventions, e.g. the state-wide effect of a change in policy. The idea of synthetic controls is to approximate one unit's counterfactual outcomes using a weighted combination of some other units' observed outcomes. The motivating question of this paper is: how does the SC strategy lead to valid causal inferences? We address this question by re-formulating the causal inference problem targeted by SC with a more fine-grained model, where we change the unit of the analysis from "large units" (e.g. states) to "small units" (e.g. individuals in states). Under this re-formulation, we derive sufficient conditions for the non-parametric causal identification of the causal effect. We highlight two implications of the reformulation: (1) it clarifies where "linearity" comes from, and how it falls naturally out of the more fine-grained and flexible model, and (2) it suggests new ways of using available data with SC methods for valid causal inference, in particular, new ways of selecting observations from which to estimate the counterfactual.

1 Introduction

Since their introduction (Abadie and Gardeazabal, 2003; Abadie et al., 2010), synthetic control (SC) methods have become commonplace for estimating causal effects with panel data.

Consider the following example. In 1988 California implemented a large-scale tobacco control program,

	—1970—	—1971—	—1972—		—1988—	—1989—
CA	123	121	123.5		90.1	82.4
AL	89.8	95.4	101.1		112.1	105.6
AK	100.3	104.1	103.9	...	121.5	118.3
VA	124.3	128.4	137		129.5	122.5
WI	106.4	105.4	108.8		102.6	100.3
⋮	⋮	⋮	⋮		⋮	⋮
WY	132.2	131.7	140.0		114.3	111.4

Table 1: We observe the annual per-capita cigarette sales in packs of different states from 1970 to 1989. The blue shaded cell reports California's sales under a tobacco program. The remaining cells report other states' sales without a tobacco program.

which increased the tobacco tax by 25 cents. Abadie et al. (2010) uses SC to study the effect of this program on the average cigarette consumption in California. The dataset contains the annual per-capita cigarette sales across a number of states, where none of the states other than California implemented a similar tobacco program. This dataset is illustrated in table 1.

We observe that smoking in California decreased after the tobacco program. However, we do not know whether the decrease is caused by the program or by other causes. Thus, to assess this difference, our goal is to estimate California's *counterfactual outcome*. What would the 1989 smoking rate of California had been if its tobacco program had *not* been implemented?

SC is a method to solve this problem. It uses the data from before 1989—when neither California nor the other states had implemented the tobacco program—to learn a model of California's smoking rate as a weighted combination of the other states' smoking rates. SC then uses the fitted weights to estimate California's counterfactual smoking rate in 1989.

In the general terminology of SC, California is the *target*, the other states are the *donors*, and the tobacco

program is the *intervention.* A typical application of SC involves aggregated time series data, such as in table 1, with one target unit and a number of donor units. These units are often "large" units, such as states (Abadie et al., 2010; Bohn et al., 2014; Cunningham and Shah, 2018), counties (Abadie and Gardeazabal, 2003), and districts (Bifulco et al., 2017). The causal question is one about the target's counterfactual, after the intervention.

To justify SC estimators, existing works often make parametric assumptions about the true data generating process (DGP) of the potential outcomes of the aggregated data. A common assumption is that the potential outcomes under control (no tobacco program) are generated according to a linear factor model with additive noise (Bai, 2009). Follow-up works develop different estimators based on this assumption (Xu, 2017; Imbens et al., 2021).

The purpose of this paper is to investigate the assumptions behind SC. When is it suitable to assume a linear factor model, and why can we write the target outcome as a weighted combination of the donors? We will show how to recover the SC methodology, but without making explicit parametric assumptions about the DGP of the potential outcomes.

The key idea behind this analysis is to construct a more fine-grained model of potential outcomes, one where the unit of analysis shifts from a "large unit" to a "small unit." In the example, this change means that the potential outcomes are defined by individuals rather than states. We assume that different states index different distributions of individuals, and the state-level outcome of interest (average smoking rate) is an average of the individual outcomes in each state.

Under this fine-grained model, the causal effect targeted by classic SC is the sub-population average treatment effect (SATE), averaged over the individuals. We will derive sufficient conditions for the non-parametric causal identification of the SATE, and we will see that it uses the SC strategy. That is, we will derive sufficient conditions for the existence of a set of donors and weights such that their weighted combination can be used to approximate the effect.

This identification result further suggests new settings in which to use the SC estimators, new ways to define the donors, and good heuristics for selecting auxiliary covariates. In more detail, there are several implications of this way of deriving the SC methodology.

First, the SC literature usually assumes that the control potential outcomes are generated according to a linear factor model. With the fine-grained model, we will see that the linear factor model form needs not be assumed a priori. Rather, the factor form is a natural consequence of invariance assumptions across groups (states) and time; and linearity is a consequence of the fact that expectation is a linear operator. A practical implication of this perspective is that SC can be used even if the individual-level DGP is non-linear.

Second, the SC literature does not generally offer guidance about the selection of the donors or how the choice of donors might affect the corresponding estimate. When reasoning from the fine-grained model, we will see how causal identifiability depends directly on properties of the target and selected donors; it may be possible to write California as a weighted combination of New York and Nevada, but not as a (non-zero) weighted combination of New York, Nevada, and Florida. Practically, the identification results suggest how to best choose donors for good SC estimation, and show that SC does not require that donors belong to the same type of group-level data. For example, we may correctly use data from Chicago (a city) to approximate the counterfactual of California (a state).

Finally, the analysis here provides a general heuristic of deciding which auxiliary covariates are suitable to include and which are not. We will see that alternative measurements of the target outcomes (e.g. per-capita tobacco spending measured in USD) are, in general, suitable auxiliary covariates, whereas summaries of group-level characteristics (e.g. average age in states) may be unsuitable auxiliary covariates.

Organization. The paper proceeds as follows. In section 1.1, we review related work. In section 2, we formulate the tobacco tax example using both classical SC and individual-level potential outcomes and introduce the corresponding estimands. In section 3.1, we introduce the SC estimators and review the common assumptions in classical SC. In section 3.2, we introduce the fine-grained model. In section 3.3, we establish sufficient conditions for causal identification of the estimand using the SC strategy. In section 3.4, we discuss implications of the reformulation and the identification result, how it points to new settings in which to use SC estimators and new ways to define donors. In section 4, we use the identification result to reason about what are suitable auxiliary covariates. In section 5, we study these implications using simulation studies and the tobacco tax dataset. Finally, in section 6, we discuss limitations and future work.

1.1 Related Work.

This paper builds on the literature pioneered by Abadie and Gardeazabal (2003); Abadie et al. (2010). A large part of the literature is about novel estimators (Abadie and Imbens, 2011; Abadie et al., 2015;

Wong, 2015; Doudchenko and Imbens, 2016; Xu, 2017; Amjad et al., 2018; Ben-Michael et al., 2018; Abadie and L'hour, 2019; Amjad et al., 2019; Arkhangelsky et al., 2019; Li, 2020; Agarwal et al., 2020; Athey et al., 2021; Imbens et al., 2021) and inference methods (Abadie et al., 2010; Doudchenko and Imbens, 2016; Ferman and Pinto, 2017; Shaikh and Toulis, 2019; Chernozhukov et al., 2021). See Abadie (2019) for an excellent review. This paper complements the existing work, as it interrogates the assumptions made by many of these estimators and inference methods.

This paper contributes to the growing effort of providing causal interpretations for synthetic controls O'Neill et al. (2016) formalize and synthesize the common assumptions in various methods for panel data inference. Bottmer et al. (2021) study SC estimators properties under a randomized experiment setup. Shi et al. (2021) develop identification and inference theory for the SC methods by drawing insights from the proximal causal inference literature (Miao et al., 2018). However, all of this existing literature performs its analysis with group-level aggregates as units. In contrast, this paper takes advantage of the nature of the group-level data and shows how SC assumptions can arise from reasoning about individual-level potential outcomes.

Finally, this paper contributes to the growing research on invariance and causality (Schölkopf et al., 2012; Bareinboim and Pearl, 2014; Peters et al., 2016; Bühlmann, 2018; Lei and Candès, 2020; Schölkopf et al., 2021). In particular, a key assumption in this paper is the independent causal mechanism principle (Peters et al., 2017).

2   Data and Problem Formulation

In this section we define the observed data, the group-level potential outcomes that underlie classical SC, and the individual-level potential outcomes that we consider in this paper. We define the causal estimand of interest in both settings. For expository purposes we omit the auxiliary covariates for now. We introduce them in section 4.

2.1   The Observed Data

We have a dataset that contains the average smoking rate of $j = 1, ..., J$ states for $t = 1, ..., T$ time periods. Let μ_{jt}^{obs} denote the average smoking rate of state j and time t. The *target* state is $j = 1$, i.e. California. It is the only state that levied tobacco taxes. The remaining states $j \geq 1$ are potential *donors*, which did not impose tobacco taxes. The intervention (the tobacco taxes) happened at time T_0. To simplify notation, we assume one post-intervention time period,

$T = T_0 + 1$, though the analysis easily generalizes to more post-intervention time periods. The number of time periods T and the potential donors J are fixed.

2.2   Classical SC Potential Outcomes

SC estimators are usually developed under the potential outcomes framework for causal inference (Splawa-Neyman et al., 1923; Rubin, 1974). Each state is considered a unit. The potential outcomes for each unit at each time period are $(\mu_{jt}(0), \mu_{jt}(1))$. These variables are the average smoking rates of state j at time t, one in the world where state j increased tobacco taxes, and one in the world where state j did not increase tobacco taxes. We assume the treatment is well-defined and no interference between the states (Rubin, 1980).

The observed outcomes are:

$$\mu_{jt}^{obs} = \begin{cases} \mu_{jt}(1) & \text{if } j = 1 \text{ and } t = T \\ \mu_{jt}(0) & \text{otherwise.} \end{cases} \quad (1)$$

In other words, we observe each states' smoking rate under no tobacco taxes *except* for California at time T, where we observe its smoking rate with the tax.

2.3   Fine-Grained Potential Outcomes

This paper considers a more fine-grained model, where we treat individuals as units and states as distributions of individuals. The pair $(Y_{ijt}(1), Y_{ijt}(0))$ denotes the potential outcomes of individual i in group j at time t, e.g. how many packs of cigarettes person i in state j consumed at time t, under increased tobacco taxes or not. Note we never make observations at an individual level, only in aggregate, but still we will reason about these variables.

We also consider the variable $X_{ijt} \in \mathbb{Z}^D$. It is a vector of causes that contribute to individual i's outcome, such as age, education level, or income. The relationship between the causes X_{ijt} and potential outcomes $(Y_{ijt}(0), Y_{ijt}(1))$ can be linear or non-linear, and can also change across time periods. We emphasize that we *will not* observe these individual-level causes, but their existence will be crucial in our reasoning about the assumptions of SC.

We assume that interventions are made at a group-level and individuals in each group comply with their group-level intervention, i.e. individuals in California do not go to Nevada to purchase tobacco and vice versa. The observed group-level averages approximate expected individual-level potential outcomes,

$$\mu_{jt}^{obs} \approx \begin{cases} \mathbb{E}\left[Y_{jt}(1)\right] & \text{if } j = 1 \text{ and } t = T \\ \mathbb{E}\left[Y_{jt}(0)\right] & \text{otherwise.} \end{cases} \quad (2)$$

The expectation is taken over the individuals i.

2.4 Causal Estimand

We are interested in the causal effect of the tobacco taxes on the average smoking of individuals in California at time period T. Using the classical SC notation, this estimand is formally defined as the unit-specific treatment effect,

$$\tau_T = \mu_{1T}(1) - \mu_{1T}(0). \tag{3}$$

Under the individual-level notation, the estimand is the sub-population average treatment effect, averaged over individuals in the target distribution,

$$\tau_T = \mathbb{E}\left[Y_{1T}(1) - Y_{1T}(0)\right]. \tag{4}$$

Remark 1. *Being precise about the causal estimand is important, because different estimands require different identifying assumptions. eq. 3 is the unit-specific treatment effect (UTE). [1] In general, UTE is incredibly difficult, if not impossible, to identify (Hernán and Robins, 2010). It involves strong parametric assumptions on the data generating mechanism and the distribution of the noise variable. In contrast, the causal estimand in eq. 4 is the an average causal effect. Causal identification of the average causal effect is easier, and in general, does not require parametric assumptions (Rosenbaum and Rubin, 1983; Imbens, 2004).*

3 A Fine-Grained Model for SC

In this section, we first review the classical SC approach to causally identify and estimate the estimand. We then develop a fine-grained model for SC, one that makes several assumptions that will eventually lead to the non-parametric identification of the causal estimand. Finally, we discuss the practical implications of reasoning about the fine-grained model, and the corresponding identification result.

3.1 Classical Synthetic Controls

The idea behind SC is to use the observed outcomes of the donor states, which did not pass a tobacco tax, to help estimate the counterfactual California outcome. Specifically, SC posits that there exists a donor set D in the potential donor pool J and a weight set $\{\beta_j\}_{j \in D}$, such that,

$$\mu_{1t}(0) = \sum_{j \in D} \beta_j \mu_{jt}(0) \quad \forall t \leq T. \tag{5}$$

The validity of SC usually relies on parametric assumptions about the control potential outcomes

$\mu_{jt}(0)$. A common assumption is that the control potential outcomes are generated according to a linear factor model plus noise (Bai, 2009),

$$\mu_{jt}(0) = \lambda_t^T \gamma_j + \epsilon_{jt}. \tag{6}$$

Here $\lambda_t \in \mathbb{R}^R$ is a time-specific vector of factors shared across different units and $\gamma_j \in \mathbb{R}^R$ are unobserved unit-specific factor loadings. The factor size R is usually assumed to be significantly smaller than the number of potential donors J and the total time periods T. The noise variable ϵ_{jt} is zero centered.[2]

Classical SC uses group-level observations to estimate the weights in eq. 5. Specifically, it fits the regularized least squares,

$$\hat{\beta} = \min_{\hat{\beta}_{j \in D}} \sum_{t=1}^{T_0} \left(\mu_{1t}^{obs} - \sum_{j \in D} \mu_{jt}^{obs} \cdot \hat{\beta}_j\right)^2 + \Upsilon(\hat{\beta}), \tag{7}$$

where $\Upsilon(\hat{\beta})$ is a prior or regularizer. SC uses the learned weights $\hat{\beta}$ to estimate the counterfactual California at time T.

To ensure a unique set of weights, existing works place restrictions on $\hat{\beta}$. For example, the original SC estimator (Abadie et al., 2010) restricts the weights to be positive and add up to one. Doudchenko and Imbens (2016) proposes an elastic-net regularizer. Robbins et al. (2017) suggests an entropy penalty.

3.2 Invariance Assumptions for the Fine-Grained Model

We now consider the fine-grained model, where we reason about an individual i, and treat each state j as a distribution of individuals. We will show how to recover the SC strategy without making an explicit parametric assumption on the data generating process of the potential outcomes.

The target estimand is the sub-population average treatment effect in eq. 4. Since the expected outcome under intervention $\mathbb{E}[Y_{1T}(1)]$ can be trivially identified, the goal is to causally identify the control expected outcome $\mathbb{E}[Y_{1T}(0)]$ from the data distribution.

In this subsection, we discuss the invariance assumptions that will lead to causal identification. In the next subsection, we will complete the derivation of how to identify the expected counterfactual.

The first assumption is that of an independent causal mechanism (ICM) (Schölkopf et al., 2012, 2021). ICM is a principle that the conditional distribution of each

---

[1]The more common name for eq. 3 is the individual treatment effect, where the units are individuals. To avoid potential confusion, we use UTE instead of ITE.

[2]The original SC (Abadie et al., 2010) for the tobacco tax example assume a variant of the factor model in eq. 6, which we discuss in section 4.

variable given its causes (i.e. its "mechanism") does not inform or influence the other conditional distributions (Schölkopf et al., 2012).

In this context, the assumption means that the causal mechanism of an individual's tobacco consumption $Y_{ijt}(0)$ is independent of the distribution of their causes X_{ijt}. If we know all the potential causes of an individual's tobacco consumption, then the distribution of the control potential outcome is independent of which distribution (i.e. state) the individual is from.

A1. (Independent Causal Mechanism)
Conditional on the causes X, the potential outcome $Y(0)$ is independent of the population distribution j. For population distribution j at time $t \leq T$, the joint distribution of $Y(0)$ and X is

$$P_{jt}(X, Y(0)) = P_{jt}(X)P_t(Y(0) \mid X). \qquad (8)$$

The assumption says that the distribution of individual causes X can vary across states and time, but the conditional outcome $Y(0) \mid X$ only varies by time. For each time point t, if we know all the potential causes of an individual's smoking behavior, which state they are from does not provide any additional information about the distribution of their control potential outcome.

Using eq. 8, we rewrite the expected counterfactual as

$$\mathbb{E}\left[Y_{jt}(0)\right] = \sum_x \underbrace{\mathbb{E}_t\left[Y(0) \mid X = x\right]}_{\lambda_t} P_{jt}(X = x), \qquad (9)$$

where $\mathbb{E}_t$ denotes the expectation with respect to $P_t(Y(0)|X = x)$, a distribution that is invariant across states.

eq. 9 is similar to the factor model in eq. 6, where the vector of factors λ_t is the vector of conditional expected outcomes, but notice we did not make any assumptions about the relationship between the causes X_{ijt} and potential outcomes $Y_{ijt}(0)$. Rather, the linearity in eq. 9 comes from the independent causal mechanism and iterated expectation — the group-level outcomes are the averages of individual-level outcomes.

The second assumption is one of stable distributions. Given the target and the *selected donors*, i.e. donors that will be used to construct the SC, we can further decompose the causes X into causes that differentiate the target and the selected donors, and causes that are invariant.

A2. (Stable Distributions) Decompose the causes into two subsets $X = \{U, S\}$. Let S denote the subset that differentiates the target from the selected donors,

i.e., its distribution in the target group is different from its distribution in the selected donor groups. We assume that, for all groups, the distribution of S does not change for all time periods $t \leq T$,

$$P_{jt}(X) = P_j(S)P_t(U \mid S). \qquad (10)$$

In other words, The subset S contains causes that vary across states, but are invariant across time. The conditional distribution of U varies by time but is invariant across states. We call S the *minimal invariant set*.

Remark 2. *The minimal invariant set S is determined by the choice of the selected donors. For example, if we choose New Zealand as a donor to California, then the set S may be the same as X, because these two groups are very different. If the donor is a "twin California," then the set S is an empty set.*

With these assumptions in hand, we use eq. 9 and eq. 10 to rewrite the expected counterfactual,

$$\mathbb{E}\left[Y_{jt}(0)\right] = \sum_s \underbrace{\mathbb{E}_t\left[Y(0) \mid S = s\right]}_{\lambda_t} \underbrace{P_j(S = s)}_{\gamma_j}. \qquad (11)$$

Comparing eq. 11 with the factor model in eq. 6, we can see that the conditional expectations are analogous to the time varying factors λ_t. The probabilities are analogous to the state specific factor loadings γ_j. What this equation shows is that with the two invariance assumptions, the expected outcome is naturally expressed as a factor model. Using the fine-grained model, we discover that the "linearity" in SC comes from aggregation, and the factor model arises from the invariance assumptions A1 & A2.

3.3 Causal Identifiability

While eq. 11 is similar to the factor model in eq. 6, it does not guarantee causal identifiability. The reason is that the cardinality of the minimal invariant set S can be very large. In particular, if the cardinality of S is larger than the number of donors then we cannot write the target as a weighted combination of the donors.

Here, we establish sufficient conditions for the causal identifiability of the causal estimand using the SC strategy. Causal identifiability is about *whether we can* express a causal estimand as a parameter of the observed distributions (Pearl, 2000). In classical SC, causal identifiability is assumed in eq. 5, that is, there exist a set of donors and weights, such that the target's counterfactual can be approximated as a weighted combination of the donors' observed outcomes.[3]

---

[3]Note, that causal identifiability is different from model identifiability, which is about whether we can recover a unique set of parameters from data.

To establish causal identifiability, we need to make two more assumptions about the data distribution.

A3. (Sufficiently Similar Donors) Let D be the set of donors used to construct the synthetic control and S be the minimal invariant set for the target and the selected donors. The donors are sufficiently similar if the cardinality of the donor set is greater than or equal to the cardinality of the minimal invariant set,

$$|D| \geq |S|. \tag{12}$$

As discussed in remark 2, the cardinality of the minimal invariant set S is determined by the selected donors. If the selected donors are very different from the target, the cardinality of S is large. If they are similar, the cardinality of S is small.

A4. (Target Donors Overlap) Let $\{s_1, ..., s_R\}$ be the support of S. There exists at least one donor distribution j in the selected donor set D, where $P_j(S = s) > 0$.

A4 implies that every type of individual living in California with characteristics $S = s$ might also live in one of the selected donor states.

A3 and A4 are two assumptions on the distribution of the minimal invariant set S. While we do not directly observe S in practice, we can still use it to reason about the differences among the population distributions, and discuss its influence on causal identifiability.

Finally, we derive sufficient conditions for non-parametric identification of the causal estimand using the SC strategy.

Theorem 1. *(Causal Identifiability) Assume the causal mechanism is independent (A1), the target and the selected donors are stable during the periods of investigation (A2), the selected donors are sufficiently similar to each other (A3), and there is overlap between the target and the selected donors (A4). Then there exists a set of weights $\{\beta_j\}_{j \in D}$, such that across all time periods, the target's counterfactual can be written as a weighted combination of the donors' outcomes,*

$$\mathbb{E}\left[Y_{1t}(0)\right] = \sum_{j \in D} \beta_j \mathbb{E}\left[Y_{jt}(0)\right] \quad \forall t \leq T. \tag{13}$$

The proof is in appendix A.

The conditions in theorem 1 lead to the assumption (eq. 5) commonly made in the literature: the existence of synthetic controls. Note that we have arrived at eq. 5 without making any parametric assumptions about the mechanism generating individual-level potential outcomes.

3.4 Implications

We have presented a set of causal assumptions that justify the SC methodology. What are the implications of these results?

Mixing types of donors. The fine-grained potential outcomes model in section 2.3 provides insights about what can be treated as a donor. Classical SC typically treats a state as a unit, and chooses donors as other states. For example, Abadie et al. (2010) excluded the District of Columbia as a possible donor. The fine-grained model implies that each donor need only be a group of individuals, and does not need to be the same type of group as the target. A donor's influence on the SC estimator has to do with how different it is to the target and other donors, i.e., the cardinality of the minimally invariant set. For example, in section 5, we will see that we can use districts or regions as potential donors to a target state.

SC with non-linear DGPs. Previous work on SC begins with a linear factor model assumption, such as the one in eq. 6 (Xu, 2017). However, it is unclear whether the role of the parametric model is for causal identification, for statistical necessity, or for notational convenience. As discussed in remark 1, because the estimand is the unit-specific treatment effect, a natural interpretation of eq. 6 is that it is an assumption on the data generating mechanism. Assuming the true mechanism is linear can be an unrealistic assumption.

In contrast, using the fine-grained model, we explained why, fundamentally, linearity can be a reasonable assumption in SC. We cast the estimand as an average effect over sub-populations, and SC as a population re-weighting algorithm. It make clear that the linear factor model in eq. 6 encodes invariance assumptions for causal identifiability, and that the linearity in eq. 6 arises from expectation being a linear operator. A practical implication of this perspective is that the causal identification result holds even if the fine-grained model involves a nonlinear mechanism. Thus SC estimators are valid in settings where the individual-level DGPs are nonlinear.

The Role of S and its relations to the donors. Classical SC assumes the latent factors λ_t in eq. 6 are fixed and low rank (Abadie, 2019). Consequentially, we may be tempted to use information from all the donors. We show that different donors can lead to different latent factors. Specifically, in eq. 11, we draw the analogy between the size of the factors and the cardinality of the minimal invariant set S. We show that the minimal invariant set S is *determined* by the donors used to construct the synthetic control. Differ-

ent donor sets can lead to different minimal invariant sets and naively including additional donors may lead to the non-existence of synthetic controls. Whether the assumption in eq. 5 holds is a property of the target and the selected donors.

4 Auxiliary Covariates

So far, we have discussed how to analyze panel data as in table 1. In many practical settings, we may find additional state-specific covariates, such as the percentage of teenagers in each state.

Previous work has integrated the auxiliary covariates into the linear factor model (Abadie et al., 2015). For example, Abadie et al. (2010) posits the following,

$$\mu_{jt}(0) = \delta_t + \lambda_t^T \gamma_j + \theta_t^T A_j + \epsilon_{jt}, \qquad (14)$$

where $\theta_t \in \mathbb{R}^R$ and $\delta_t \in \mathbb{R}^R$ are time-specific vectors of factors shared across different units and $A_j \in \mathbb{R}^k$ is a vector of K observed state-specific covariates.

Abadie et al. (2010) assumes that there exists a set of weights such that for all $k \leq K$ and $t \leq T$,

$$A_{1k} = \sum_{j \in D} \beta_j A_{jk} \quad \text{and} \quad \mu_{1t}(0) = \sum_{j \in D} \beta_j \mu_{jt}(0). \quad (15)$$

As a result, one way to use auxiliary covariates is to analyze them in parallel with the outcomes to compute the SC weights. (Abadie et al., 2010, 2015; Ben-Michael et al., 2018; Botosaru and Ferman, 2019).

This use of auxiliary covariates requires us to assume that the underlying data generating process is linear. A natural question is whether we can still use auxiliary covariates to construct the SC weights without assuming the underlying DGP is linear. Here, we will use the fine-grained model to reason about *suitable auxiliary covariates*, those that inherently satisfy eq. 15.

Using eq. 11, we can reason about two types of suitable auxiliary covariates. The first type are the state-specific probabilities for the variables in the minimal invariant set S: $P_j(S = s)$. We observe that in eq. 11, the relationships between the probabilities and the outcomes are linear. Therefore, the probabilities have the same relationship with one another that the outcomes have with each other. For example, suppose the variable "age" differentiates target and donor distributions, i.e., "age" is in S, then "percentage of young adults" would be a suitable covariate.

Note that some group-level summaries of individual-level characteristics may not be suitable auxiliary covariates. For example, consider the average age (within a state). Since we *do not* assume a linear relationship between the individual-level characteristics and the individual-level outcomes, we can not expect linear relationships between the group-level summaries and the group-level outcomes.

The second type of suitable auxiliary covariates are different measurements of the target outcomes. With a bit of algebra, we can see that the SC weights in eq. 13 are combinations of the state-specific probabilities on the minimal invariant set: $P_j(S = s)$. Recall that the minimal invariant set S is solely determined by the target and the selected donor and that they are not influenced by the outcome measurements. Therefore, the SC weights should be invariant to different measurements of the outcome variable Y. For example, per-capita tobacco spending measured in USD would be a suitable auxiliary covariate to the target outcome: average cigarette consumption measured in packs.

5 Empirical Studies

We use simulations and tobacco tax data to study this perspective on SC and the implications of the theory around the fine-grained model.[4] We find the following.

1. There is no reason the donors need to come from the same type of group as the target. We can still recover meaningful causal estimates when using donors from a different type.

2. The SC estimators are valid even when the individual-level causal mechanisms are nonlinear.

3. The cardinality of the minimal invariant set S is dependent on the donor and target choices and is crucial to whether the weights can be generalized to the counterfactual.

4. We may improve SC estimates when we include suitable auxiliary covariates in the analysis. However, unsuitable covariates will bias the estimates.

Simulations. Following section 3, we first generate individual-level data, then construct group-level summaries. The individual-level covariates X take $K = 12$ values. Individual-level outcomes are derived from a set of non-linear and time-varying functions. We create one target group and 5 donor groups. Each group has a different composition of individuals. The compositions do not change over time.

We consider T time periods. For each group, at each time period, we sample 2000 individuals according to its population composition. The group-level summaries, μ_{jt}^{obs}, are the average outcomes of individuals in group j at time point t. The randomness is on an individual level, instead of a group level.[5]

---
[4]code: github.com/claudiashi57/fine-grained-SC
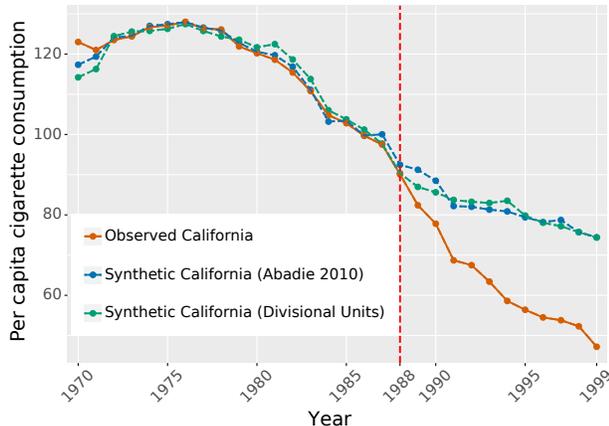[5]More simulation details are in appendix B

Figure 1: Donors are not required to be of the same type: we use a linear combination of divisions to capture the outcome trends of California (a state) with high fidelity.

Figure 2: SC estimators remain valid when causal mechanisms generating individual-level outcomes are nonlinear. The figure shows the observed and counterfactual prediction losses when the group-level measurements are the "mean" or the "median".

We create two knobs in the simulation, S and T. The variable T denotes the number of time periods, corresponding to the number of data points when fitting and evaluating the SC estimator. The set S is the minimal invariant set. We use $|S|$ to denote the cardinality of S, and specifically how much the target and selected donor distributions differ from each other. $|S|$ ranges from 0 to K. If the donors are identical to the target $|S| = 0$. If the donors are very different in all aspects of the population composition $|S| = K$. We do not observe the minimal invariant set S and its cardinality. The SC estimators do not use S or $|S|$.

Prop 99. Following Abadie et al. (2010), we analyze data about the Prop 99 tobacco program. We use the state-level data for the period 1970–2000. We exclude states that also implemented large tobacco programs during the time frame and the states that raised tobacco tax by more than 50 cents, resulting in 39 potential donors. The outcome measurement is the per-capita cigarette sales in packs. The main distinctions to Abadie et al. (2010) are that (1) we include Washington DC in the donor pool, and (2) exclude the state-level covariates.

Methods and evaluation. For the Prop 99 example, we use the original SC estimator, with positive weights that sum up to one. Since we cannot observe counterfactuals, we evaluate the estimation quality by plotting out the estimated counterfactuals.

For the simulation studies, we use ordinary least squares (OLS) as the SC estimator. We use 75% of the data to fit the estimator and 25% of the data for out-of-sample evaluation. The evaluation metric is the mean squared error averaged over data (time) points.
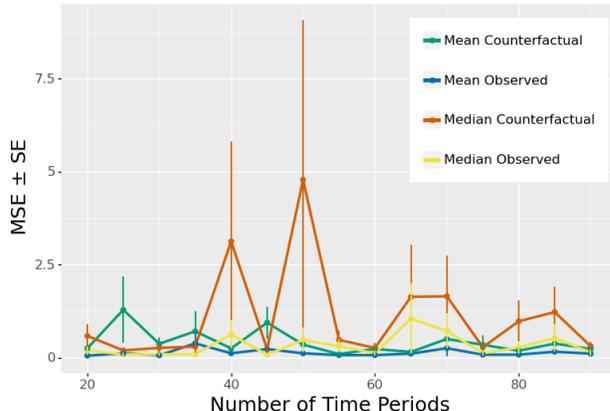
For expository purposes, when discussing the estimation quality, we use "observed" and "counterfactual" instead of "in-sample" and "out-of-sample".

(1) Mixing types of donors. As discussed in section 3.4, there is no reason to restrict donors to be of the same type as the target. To study this possibility empirically, we consider a type of donor that is different from states. Since 1950, the United States Census Bureau has defined nine statistical divisions based on geographical location, e.g. New England, Mountain. Each division contains several states. We construct divisional-level donors using the United States census of 1990. The average smoking rate of each division is a weighted combination of the smoking rate in its corresponding states, weighted by their population.

We use the original SC estimator to construct a synthetic California using these divisional-level data. We compare the counterfactual prediction of the divisional-level SC estimator with the original SC estimator. As shown in fig. 1, the synthetic California constructed by divisional-level data can capture the outcome trends of California with high-fidelity. We interpret the weights of the SC estimator in appendix C.

(2) SC with no-linear DGPs. section 3.2 argues that the linearity in SC comes from aggregation, rather than a linear individual-level data generating process. We study this claim empirically using the nonlinear data simulation. We have data from six groups, one is the target and the other groups are the donors. We set the cardinality of S to 5 and consider a range of time periods for the data, from $T = 20$ to $T = 90$.

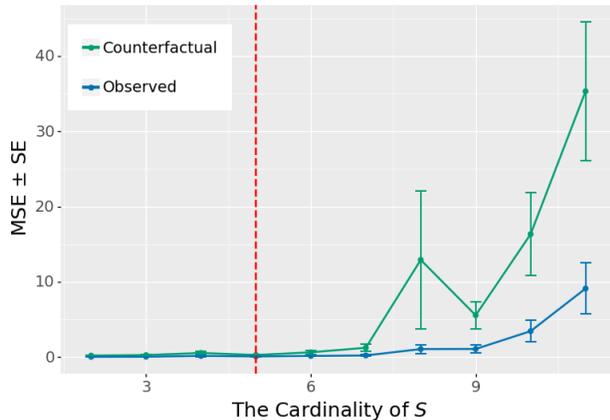For a given number of time periods, we construct two

Figure 3: The more different the donors are, the worse the SC counterfactual estimates become. The red dotted line denotes the number of donors used for constructing the synthetic controls.

panel datasets. One contains the *average* outcomes of individuals in groups. The other includes the *median* outcomes of individuals in groups. Note that the "mean" is linear, but the "median" is nonlinear.

We apply the SC estimator to both panel datasets and evaluate the predictive performance on the observed and counterfactual data. As shown in fig. 2, when the individual-level DGP is nonlinear, the SC estimator can still produce valid counterfactual estimates for the average outcomes. Of course, linearity does not come for free. SC estimates are valid in this example because the "mean" is a linear function, and the cardinality of the set S is not greater than the number of available donors. In contrast, when the measurement is the "median", the SC estimator fails at predicting the counterfactual, because the median is a nonlinear function. Thus, we cannot expect a linear relationship between the median of the target group and the median of the donor groups.

(3) The minimal invariant set S is critical to causal identification. As discussed in section 3.3, the minimal invariant set S and whether eq. 5 holds are properties of the target and the donors. If we choose donors that are drastically different from the target, the SC weights learned with the observed data may not generalize to the counterfactual data. We use nonlinear simulations to study the relationship between the donor choice and the quality of the counterfactual estimates. We fix the number of the donors to 5, the number of time periods to 20, and increase the cardinality of S from 2 to 11.

As shown in Figure 3, once the cardinality of S surpasses the number of available donors, the counter-

factual estimation error increases significantly. Importantly, it is increasing at a significantly faster rate than the observed error. We cannot determine whether the SC estimator can produce valid counterfactual estimates from the observed dataset alone.

Table 2: Using suitable covariates improves the estimation quality, whereas unsuitable covariates hurt the estimation quality. The table reports the mean squared error and standard error over 100 simulations.

MSE ± SE	Observed	Counterfactual
Outcome Only	.07 ± .03	.14 ± .05
Suitable Covariates	.06 ± .02	.13 ± .05
Unsuitable Covariates	.06 ± .02	.24 ± .13

(4) Auxiliary Covariates. Finally, we study what happens to the counterfactual estimates when we include auxiliary covariates. Using the simulations, we construct 10 suitable auxiliary covariates and 10 unsuitable auxiliary covariates. The suitable covariates are the averages of the sine transformation of individual-level outcomes. The unsuitable covariates are averages of the individual-level covariates. We fix the number of the donors to 5, the cardinality of the minimal invariant set to 5, and the number of time periods to 15. We examine the observed and counterfactual estimation quality when including suitable covariates and including unsuitable covariates. As shown in table 2, incorporating suitable covariates may improve the counterfactual estimation, whereas including unsuitable covariates hurts the counterfactual estimation. Notably, we may not infer whether an auxiliary covariate is suitable by looking only at the SC estimator's fit to the observed data.

6 Discussion & Future Work

In this paper, we develop a fined-grained model for synthetic controls. Using the tobacco example, we show that the "linearity" in SC comes from aggregation. We further establish sufficient conditions for the non-parametric identification of the causal estimand and discuss several practical implications.

While this paper points to new ways of applying SC methods, the validity rests on the strong assumptions that an analyst must carefully consider. For future work, we plan to establish uncertainty quantification methods that are compatible with the fine-grained model. We also plan to develop sensitivity analyses that show how violations of the assumptions change the estimated effects.

References

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

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*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, 93(1):113–132.

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1):1–11.

Abadie, A. and L'hour, J. (2019). A penalized synthetic control estimator for disaggregated data. *Preliminary and incomplete*.

Agarwal, A., Alomar, A., Cosson, R., Shah, D., and Shen, D. (2020). Synthetic interventions. *arXiv:2006.07691*.

Amjad, M., Misra, V., Shah, D., and Shen, D. (2019). mrsc: Multi-dimensional robust synthetic control. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 3(2):1–27.

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. (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, pages 1–41.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.

Bareinboim, E. and Pearl, J. (2014). Transportability from multiple environments with limited experiments: Completeness results. In *Advances in neural information processing systems*.

Ben-Michael, E., Feller, A., and Rothstein, J. (2018). The augmented synthetic control method. *arXiv:1811.04170*.

Bifulco, R., Rubenstein, R., and Sohn, H. (2017). Using synthetic controls to evaluate the effect of unique interventions: The case of say yes to education. *Evaluation review*, 41(6):593–619.

Bohn, S., Lofstrom, M., and Raphael, S. (2014). Did the 2007 legal arizona workers act reduce the state's unauthorized immigrant population? *Review of Economics and Statistics*, 96(2):258–269.

Botosaru, I. and Ferman, B. (2019). On the role of covariates in the synthetic control method. *The Econometrics Journal*, 22(2):117–130.

Bottmer, L., Imbens, G., Spiess, J., and Warnick, M. (2021). A design-based perspective on synthetic control methods. *arXiv:2101.09398*.

Bühlmann, P. (2018). Invariance, causality and robustness. *arXiv:1812.08233*.

Chernozhukov, V., Wüthrich, K., and Zhu, Y. (2021). An exact and robust conformal inference method for counterfactual and synthetic controls. *Journal of the American Statistical Association*, pages 1–44.

Cunningham, S. and Shah, M. (2018). Decriminalizing indoor prostitution: Implications for sexual violence and public health. *The Review of Economic Studies*, 85(3):1683–1715.

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.

Hernán, M. A. and Robins, J. M. (2010). Causal inference.

Imbens, G., Kallus, N., and Mao, X. (2021). Controlling for unmeasured confounding in panel data using minimal bridge functions: From two-way fixed effects to factor models. *arXiv:2108.03849*.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and statistics*, 86(1):4–29.

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

Miao, W., Geng, Z., and Tchetgen Tchetgen, E. J. (2018). Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*.

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., and Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, 16(1):1–21.

Pearl, J. (2000). *Causality: models, reasoning and inference*. Springer.

Peters, J., Bühlmann, P., and Meinshausen, N. (2016). Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*.

Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference: foundations and learning algorithms*. The MIT Press.

Robbins, M. W., Saunders, J., and Kilmer, B. (2017). A framework for synthetic control methods with high-dimensional, micro-level data: evaluating a neighborhood-specific crime intervention. *Journal of the American Statistical Association*, 112(517):109–126.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*.

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

Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, 75(371):591–593.

Schölkopf, B., Janzing, D., Peters, J., Sgouritsa, E., Zhang, K., and Mooij, J. (2012). On causal and anticausal learning. *arXiv:1206.6471*.

Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., and Bengio, Y. (2021). Toward causal representation learning. *Proceedings of the IEEE*.

Shaikh, A. and Toulis, P. (2019). Randomization tests in observational studies with staggered adoption of treatment. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2019-144).

Shi, X., Miao, W., Hu, M., and Tchetgen, E. T. (2021). On proximal causal inference with synthetic controls. *arXiv:2108.13935*.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. (1990(1923)). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, pages 465–472.

Wong, L. (2015). Synthetic control estimation with individual-level data. In *Three Essays in Causal Inference*. Stanford University.

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

Supplementary Material:
On the Assumptions of Synthetic Control Methods

A   Proof For Thm 1

Theorem 1. *Assume the target and the selected donors are stable during the periods of investigation (A2), the donors are sufficiently similar to each other (A3), and there is overlap between the target and the donors (A4). Then there exists a set of weights $\{\beta_d\}_{d \in D}$, such that across all time periods, the target's counterfactual can be written as a weighted combination of the donors' outcomes,*

$$\mathbb{E}\left[Y_{1t}(0)\right] = \sum_{d \in D} \beta_d \mathbb{E}\left[Y_{dt}(0)\right] \quad \forall t \leq T. \tag{16}$$

Proof. We first show that for a fixed time point t, there exist a set of weights such that the target can be written as a weighted combination of the donors. We then show that there exists a set of weights that is invariant across time periods.

Fixing a time point t, we can write out eq. 11, $\mathbb{E}\left[Y_{jt}(0)\right] = \sum_s \mathbb{E}_t Y(0) \,|\, S = s P_j(S = s)$, for the target $j = 1$ and the selected donors D. Conceptually, we can think of the set expanded expectations as a system of linear equations, where the conditional expectations are the unknowns, and the probabilities are the scalars. A4 says that the unknowns in the target equation are also in at least one of the donor equations. A3 says that there are at least as many independent equations as the number of unknowns. Combining A3 & A4, we can "solve" the unknowns. Consequentially, we can write the target as a weighted combination of the donors, where the weights are functions of the probabilities.

A2 implies the probabilities are the same across time periods. Since the weights are functions of the probabilities, and the probabilities are invariant across time periods, the weights are also invariant across time periods.   □

B   Simulation Details.

We discuss the details of the simulation studies. Recall, the individual-level covariates take $K = 12$ values. There are 6 groups. Each group has a different composition of individuals, denoted by $\boldsymbol{P}_j$. The sparsity of the probabilities is determined by the cardinality of the minimal invariant set S.

Group-level parameters

$$\alpha_{sk} \sim Bin(1, 1 - \frac{|S|}{K})$$
$$\boldsymbol{\alpha}_s = (\alpha_{s1}, ..., \alpha_{sk})$$
$$\boldsymbol{P}_j \sim Dir(K, \boldsymbol{\alpha}_s)$$

Individual-level data

$$X_{ijt} \sim Cat(K, \boldsymbol{P}_j)$$
$$Y_{ijt} = f^t(X_{ijt}) + N(0, 1)$$
$$\mu_{jt} = \frac{1}{N_j} \sum Y_{ijt}$$

Individual-level outcomes are derived from a set of non-linear and time-varying functions. The exact functions are in the supplementary material. The individual-level covariates take $K = 12$ values. At each point t and for each group j, we sample $N_j = 1000$ individuals. The group-level summaries, μ_{jt}^{obs}, are the average outcomes of individuals in group j at time point t.

C   Prop 99 Experiment Details.

In fig. 1, we show that the synthetic California constructed by divisional level donors can capture the outcome trends of California with high fidelity. Here we describe how the donors are constructed.

Since 1950, the United States Census Bureau divided the United States into nine divisions: New England, Mid-Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific. Each division consists of several states. Following Abadie et al. (2010), we exclude states that also implemented large tobacco programs during the time frame and the states that raised tobacco tax by more than 50 cents: Massachusetts, Arizona, Oregon, Florida, Alaska, Hawaii, Maryland, Michigan, New Jersey, New York, and Washington.

We construct divisional-level donors according to the United States census of 1990. Each divisional donor is a weighted (by population) combination of its corresponding states that were not excluded for implementing similar policies. We drop the Pacific division because all states in the division have been excluded. We use the original SC estimator to construct a synthetic California using these divisional-level data.

Division	Weight
New England	.05
Mid-Atlantic	0
East North Central	0
West North Central	0
South Atlantic	0
East South Central	0
West South Central	0
Mountain	.95
Pacific	-

Table 3: Divisional-level weights used to construct the synthetic California.

As shown in table 3, the synthetic California uses only two donors: Mountain division and New England division. This is not surprising because California is geographically close to the Mountain division. Table 3 is also consistent with the result in Abadie et al. (2010), where the chosen state-level donors come from either the Mountain division or the New England division.

Changing the unit of analysis does not necessarily harm interpretability. California is a state with a high population density. According to the 1990 census, the population in California is around 30 million. The total population in the Mountain division used to construct the control is around 10 million. In contrast, the population in Montana, a donor used in Abadie et al. (2010) with 0.2 weights, is only around 0.8 million. Taking population level into consideration, it is as interpretable, if not more, to treat divisions as potential donors to California.