WORKING PAPER

Imperfect Synthetic Controls: Did the Massachusetts Health Care Reform Save Lives?

David Powell

RAND Labor & Population

WR-1246
May 2018
This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the RAND Labor and Population Unit.

RAND CORPORATION

For more information on this publication, visit www.rand.org/pubs/working_papers/WR1246.html

Published by the RAND Corporation, Santa Monica, Calif.

Imperfect Synthetic Controls:
Did the Massachusetts Health Care Reform Save Lives?*

David Powell
RAND†

May 2018

Abstract

The synthetic control method has become a valuable and widely-used technique to estimate causal effects even when more traditional fixed effects methods are inappropriate. This paper relaxes two critical assumptions required to implement the synthetic control estimator. First, the synthetic control estimator assumes that the outcomes of the treated unit are within the "convex hull" of the outcomes of the untreated units. In this paper, I show that estimation of the policy effect is possible when the treated unit composes part of the synthetic control for any of the untreated units, permitting the treated unit to be outside the convex hull of the other units. Instead of constructing a synthetic control only for the treated unit, this paper recommends creating a synthetic control for every unit. The difference in the post-treatment outcomes for each unit compared to its synthetic control is related to the corresponding difference in the policy variable, identifying the policy effect. Second, the synthetic control estimator assumes the existence of a "perfect" synthetic control, which only occurs if the outcome variable is not subject to transitory shocks. In this paper, I suggest a straightforward two-step approach which first generates predicted values of the outcome variables for each unit and uses these predicted values instead of the actual values of the outcome variable when constructing the synthetic control units. Together, these two modifications significantly reduce the restrictions imposed by the synthetic control estimator and provide asymptotically unbiased estimates of the policy effect. Simulations show that this approach outperforms the traditional synthetic control estimator. I apply the new estimator to study the mortality effects of the 2006 Massachusetts Health Care Reform and estimate that the reform reduced the mortality rate by 3%.

Keywords: synthetic control estimation, fixed effects, difference-in-differences, parallel trends, health insurance, mortality, health care reform
JEL classification: I13, C23, I18

*I received helpful feedback from participants at the Annual Health Econometrics Workshop and the Annual Health Economics Conference, especially from my discussants Juan Pantano and Colin Cameron.
†Contact information: dpowell@rand.org

1 Introduction

Empirical research often relies on panel data to compare outcome changes in a unit adopting a policy of interest to units that did not adopt the policy. It is common to assume that the adopting unit would have experienced similar changes in the outcome as the average of the non-adopting units. This assumption motivates the use of a difference-in-differences framework and estimating a specification with additive unit and time fixed effects. The synthetic control method (SCM) developed in Abadie and Gardeazabal (2003), Abadie et al. (2010), and Abadie et al. (2015) relaxes the "parallel trends" assumption and permits the creation of an optimal synthetic control, a weighted average of the control units which best approximates the treated unit in the pre-intervention period. While an unweighted average of other units may not be an appropriate control, a subset of units with non-uniform weights may provide a proper counterfactual. A major benefit of this approach is that it allows for estimation of a specification with interactive fixed effects which nests more traditional additive fixed effects models.

The synthetic control estimator uses a data-driven method to create a weighted average of the other units to act as a counterfactual and has been called "[a]rguably the most important innovation in the evaluation literature in the last fifteen years" (Athey and Imbens (2017)). This approach, however, relies on two significant restrictions. First, the method assumes that there exists a weighted average of the pre-intervention outcomes of the control units equal to the pre-intervention outcomes of the treated unit (i.e., that the treated unit belongs to the "convex hull" of the control units). Second, the synthetic control method assumes that there exists a synthetic control such that the pre-intervention fit between the synthetic control and the treated unit is *perfect*. While rarely (if ever) holding in practice, this assumption is necessary so that the pre-intervention outcomes of the control units perfectly represent the interactive fixed effects without error. This point is highlighted in Ferman and Pinto (2016), which shows that the synthetic control weights are biased if there does not exist a perfect synthetic control. This bias occurs since the outcome is also a function of transitory shocks (i.e., the specification includes an error term with positive variance).

Many empirical applications have benefited from the synthetic control method.[1] The modifications introduced in this paper are simple to implement and substantially relax the underlying assumptions necessary to implement the synthetic control method, which are likely violated in applied settings. The modified approach uses additional information by developing synthetic controls for all units, which provides benefits even when the SCM restrictions hold. I include results from several sets of simulations and find that the modified approach substantially reduces the bias associated with the SCM.

The modified approach relies on two insights. First, the SCM assumes that the pre-intervention treated unit outcomes are within the convex hull of the control unit outcomes. This paper exploits the insight that even when this convex hull assumption does not hold for the treated unit, the treated unit may be part of an appropriate synthetic control for one or more untreated units. The post-treatment outcome differences for these units are informative of the policy effect. For example, if the treated unit is given a weight of 0.4 as part of the synthetic control for unit i, then the difference in unit i's outcome and its synthetic control is, on average, equal to -0.4 times the treatment effect in the post-period. The estimator in this paper constructs synthetic controls for *all* units and estimates the policy effect by using the differences in the outcomes and the differences in the policy variable for all units. This approach provides an opportunity to place more weight on units with better synthetic control fits.

Second, the SCM assumes the existence of a synthetic control with a perfect fit in the pre-period. The motivation of the SCM is to account for an unknown number of interactive fixed effects, but these terms are unobserved. Instead, we observe outcomes, which are noisy measures of the interactive fixed effects. This noise, much like measurement error in an explanatory variable using ordinary least squares (OLS), leads to inconsistent estimates of the synthetic control weights. In this paper, I show that it is possible to estimate consistent synthetic control weights by first predicting the values of the outcome variable using flexible state-specific trends. The synthetic controls weights can be then generated using these predicted values, subsequently permitting proper estimation of the policy effect.

[1] See Kleven et al. (2013); Dustmann et al. (2017); Akcigit et al. (2016); Acemoglu et al. (2016); Billmeier and Nannicini (2013); Donohue et al. (2015); Pinotti (2015); Bohn et al. (2014); Nonnemaker et al. (2011); Hinrichs (2012); Cavallo et al. (2013); Saunders et al. (2014); Ando (2015); Bauhoff (2014); Fletcher et al. (2014); Cunningham and Shah (2014); Mideksa (2013); Nannicini and Billmeier (2011); Torsvik and Vaage (2014); Munasib and Rickman (2015); Munasib and Guettabi (2013); Eren and Ozbeklik (2016); Kreif et al. (2015); Montalvo (2011); Hackmann et al. (2015); Adhikari and Alm (2016); Adhikari et al. (2016); Zhang et al. (2016); Quast and Gonzalez (2016); Grier and Maynard (2016); Dillender et al. (2016) for a small subset of examples.

Together, these two modifications substantially improve the performance of the synthetic control approach. Both modifications deal with the non-existence of perfect synthetic controls. For this reason, I refer to the modified approach of this paper as the Imperfect Synthetic Control Method (ISCM). The lack of perfect synthetic controls is the norm in empirical applications so the proposed method should have broad applicability. I apply this modified approach to estimate the mortality effects of a comprehensive health care reform. In 2006, Massachusetts passed comprehensive health care reform. The law included an expansion of Medicaid, subsidies for low-income households, and an individual mandate. The Massachusetts reform served as a model for the Affordable Care Act (ACA) so its effects on population health are of particular interest. The Massachusetts health care reform has been studied extensively to understand its effect on an array of outcomes. Previous research has found that the health care reform improved health insurance coverage (Long et al. (2012a,b); Kolstad and Kowalski (2012); Long (2008); Pande et al. (2011)), access to care (Long (2008); Pande et al. (2011)), and use of preventive care (Miller (2012)).

The health effects of the reform have also been analyzed. Courtemanche and Zapata (2014) finds improvements in physical and mental health on multiple dimensions while Long et al. (2012a) reports positive effects on self-assessed health. Most relevant to this paper, Sommers et al. (2014) estimates that the Massachusetts health care reform was associated with a reduction in mortality. Typically, studies analyzing the Massachusetts health care reform rely on a difference-in-difference design, assuming that the national average represents an appropriate control for Massachusetts during this time period. Some studies select on states that are geographically close. Sommers et al. (2014) employs a difference-in-differences design while selecting controls using a propensity score approach. A propensity score is estimated at the county-level to select control units based on pre-2006 values of economic and demographic characteristics, such as age distribution, sex, race/ethnicity, poverty rate, median income, unemployment, uninsured rate, and baseline annual mortality. They select the top quartile of counties with the the highest propensity scores to use as a control group.

The SCM is well-suited to estimate the effects of the Massachusetts reform. Massachusetts mortality rates are not well-approximated by the national average before 2006 so constructing a weighted average which has similar mortality rates and trends prior to the implementation of the reform should improve our ability to estimate the policy effect associated with the health care reform. Using the modified approach, I estimate that the Massachusetts reform reduced mortality by 3%.

In the next section, I include more background about the synthetic control method. Section 3 discusses the two limitations of the SCM and how the proposed modifications relax these restrictions. Section 4 introduces the ISCM estimator and discusses its properties. In Section 5, I study the mortality effects of the Massachusetts health care reform while Section 6 concludes. I include simulation results throughout the paper as I introduce modifications to the SCM.

2 Background

2.1 The Synthetic Control Method

This paper builds on the synthetic control estimation technique discussed in Abadie and Gardeazabal (2003), Abadie et al. (2010), and Abadie et al. (2015). In the synthetic control framework, there are N units and T time periods. Unit 1 is exposed to the treatment in periods $T_0 + 1$ to T and unexposed in periods 1 to T_0. All other units are unexposed in all time periods. Outcomes are defined by a factor model:

$$Y_{it} = \alpha_{it} D_{it} + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \epsilon_{it}, \tag{1}$$

$$\text{where} \quad D_{it} = \begin{cases} 1 \text{ if } i = 1 \text{ and } t > T_0, \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

D_{it} represents the treatment variable and is equal to 1 for state $i = 1$ and time periods $t > T_0$, 0 otherwise. $\boldsymbol{\lambda}_t$ is a $1 \times F$ vector of common unobserved factors and $\boldsymbol{\mu}_i$ is an $F \times 1$ vector of factor loadings.[2] Gobillon and Magnac (2016) compares the synthetic control approach to panel data models with interactive fixed effects (e.g., Bai (2009)). Both approaches are useful for relaxing the restriction that the fixed effects are additive, but Gobillon and Magnac (2016) provides simulation results in which the SCM performs poorly – even relative to difference-in-differences – as the convex hull assumption is violated.

Unit 1 is the treated unit while all other units are part of the "donor pool," the units that may potentially compose the synthetic control. The purpose of the approach, as

[2] Abadie et al. (2010) also includes time fixed effects in the specification, though these can be nested in the $\boldsymbol{\lambda}_t \boldsymbol{\mu}_i$ term.

modeled in Abadie et al. (2010),[3] is to find a weighted combination of units (which exists by assumption) in the donor pool such that

$$\sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_{jt} = \boldsymbol{\mu}_{1t}, \tag{3}$$

where w_j^1 represents the weight on unit j to construct the synthetic control representing unit 1. For the traditional SCM, it is unnecessary to designate that the weight is *for* the construction of the synthetic control for unit 1, but this notation will be useful below. If this condition holds, then $\sum_{j=2}^{N} w_j^1 Y_{jt}$ provides an unbiased estimate of Y_{1t} for $D_{1t} = 0$. After policy adoption ($t > T_0$), $\sum_{j=2}^{N} w_j^1 Y_{jt}$ serves as a counterfactual for the treated unit.

The synthetic control weights are constrained to be non-negative and to sum to one. They are generated to minimize the difference between the pre-treated outcomes of the treated unit and the unit's synthetic control. I assume that there are no other control variables. In the setup in Abadie and Gardeazabal (2003) and Abadie et al. (2010), additional covariates are also included additively in the main equation (represented above by equation (1)). In this paper, I assume that all pre-treatment outcomes (and only pre-treatment outcomes) are used to generate the synthetic control weights. This is common in applications using the SCM (see Cavallo et al. (2013) for one example). Kaul et al. (2015) notes that the synthetic control estimator does not use the covariates when each pre-treatment outcome is used to generate the synthetic control weights. Due to concerns of overfitting, it is often recommended that covariates and a limited set of pre-treatment outcomes be used to create the synthetic control weights. In this paper, I assume large T such that overfitting is less likely to be a concern. Moreover, the modification of this paper requires a two-step process such that the pre-intervention outcome values are not used but, instead, predicted values of this outcome, further reducing concerns of overfitting. To facilitate comparisons between the SCM and the modifications in this paper, I assume no covariates in discussing both estimators. I include a simple extension in Section 4.6.2 to account for time-varying covariates.

Let $\mathbf{w}_1 = (w_2^1, \ldots, w_N^1)'$ such that $w_j^1 \geq 0$ for all $2 \leq j \leq N$ and $\sum_{j=2}^{N} w_j^1 = 1$. The

[3] Equation (3) differs from the assumption specified in Abadie et al. (2010) which replaces the $\boldsymbol{\mu}$ terms with Y. This difference is discussed more in Section 3.2 but is not relevant to the discussion here.

weights are estimated using

$$\hat{\mathbf{w}}_1 = \underset{\boldsymbol{\phi}_1}{\operatorname{argmin}} \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{N} \phi_j^1 Y_{jt}\right)^2$$
$$\text{s.t.} \quad \hat{w}_j^1 \geq 0 \quad \text{for all} \quad 2 \leq j \leq N \quad \text{and} \quad \sum_{j=2}^{N} \hat{w}_j^1 = 1.$$

In words, the synthetic control estimator involves a constrained optimization to find the weighted average of the donor states which is "closest" to the treated unit in terms of pre-treated outcomes. Because I assume that all pre-treatment outcomes are used to create the synthetic control weights, I do not consider weighting matrices. The SCM uses a two-step process in which some variables included in the minimization are weighted more than others if they are better predictors of the pre-intervention outcomes. Since I use the pre-intervention outcomes themselves, this weighting matrix is unnecessary. The treatment effect estimate for period $t > T_0$ is

$$\hat{\alpha}_{it} = Y_{it} - \sum_{j=2}^{N} \hat{w}_j^1 Y_{jt}. \tag{4}$$

The SCM permits estimation of a treatment effect for each post-treatment time period, though it is common to report an aggregate post-treatment effect.

The SCM has proven valuable in applied work as a means of relaxing the parallel trends assumption required by additive fixed effect models. The framework of equation (1) nests the more typical additive fixed effect model which assumes

$$\boldsymbol{\lambda}_t = \begin{bmatrix} 1 & \phi_t \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu}_i = \begin{bmatrix} \gamma_i \\ 1 \end{bmatrix}, \tag{5}$$

such that $\boldsymbol{\lambda}_t \boldsymbol{\mu}_i = \gamma_i + \phi_t$. Difference-in-differences is a workhorse of applied microeconomics, and the additive fixed effects model is still available in this framework. The SCM permits additional flexibility, allowing for rich and flexible state-year shocks to be correlated with treatment assignment. In a typical setup in which a state adopts a policy, it is impossible to account for state-time interactions. Equation (1), however, allows for underlying shocks and trends which vary at the state-time level while nesting the more traditional additive fixed effects models.

2.2 Modified Synthetic Control Approaches

Recent working papers have also addressed how to modify the SCM when the convex hull assumption does not hold. Doudchenko and Imbens (2016) introduce a similar method to the SCM which does not enforce the convex hull assumption by (1) permitting negative weights; (2) allowing the sum of the weights to not equal one; and (3) permitting permanent additive differences across units. The first two modifications necessitate regularizing the choice of weights. The SCM restrictions often identify unique weights so relaxing these restrictions requires use of a penalty term. In contrast, ISCM maintains the SCM restrictions, aiding the choice of unique weights without regularization, but permits identification of the treatment effect by creating synthetic controls for non-treated units.

The third modification can extend synthetic control estimation to cases in which the convex hull assumption does not hold simply because of level differences across units. Including an additive fixed effect for each unit by itself does not necessarily account for situations in which the convex hull restriction does not hold. It will adjust for cases in which the treated unit generally has the largest or smallest outcomes throughout the pre-period. However, the convex hull assumption may not hold even after such level differences are adjusted for. If an additive term is the only reason that the convex hull assumption does not hold, then ISCM should address this problem without explicitly including an additive unit term since the interactive effects in equation (1) permit (and nest) additive fixed effects.[4] Note further that if it is independently desirable to include an additive unit fixed effect term in equation (1) for an application, ISCM can be extended in the same manner that Doudchenko and Imbens (2016) extend the SCM to include additive differences.

Athey et al. (2017) introduces a matrix completion approach which the authors state converges to a synthetic control-type estimator when N is small and T is large. The benefits of this method (as it converges to the SCM) are similar to those discussed for Doudchenko and Imbens (2016) since the matrix completion method places no restriction on the implicit "weights" and instead relies on the inclusion of a penalty term in the optimization. It is also important to highlight that the approaches in Doudchenko and Imbens (2016) and Athey et al. (2017) use the outcomes to create the controls for each unit which, as discussed below, is potentially problematic. ISCM uses a two-step approach which first predicts values of the

[4] The additive fixed effects are only problematic if they are the reason that the convex hull assumption does not hold. But ISCM provides an approach to estimate the policy effect even when the convex hull assumption does not hold.

outcomes and uses these predicted outcomes.

2.3 Massachusetts Health Care Reform

The effects of the Massachusetts Health Care Reform are of general interest, but may be difficult to study in a difference-in-differences framework. The focus of this paper is understanding the mortality effects of the reform and, as described in more detail in Section 5, my primary outcome of interest is deaths per 100,000 people where deaths are defined as those which are "amenable to health care" (see Sommers et al. (2014)). To test for the appropriateness of a traditional additive fixed effects model, I estimate an event study. I regress the mortality rate on time (month) fixed effects, state fixed effects, and Massachusetts-year indicators. I normalize the 2005 Massachusetts effect to zero and present the point estimates in Figure 1.

Figure 1: OLS Event Study Estimates

Notes: I estimate the policy effect for each year, both before and after enactment in an event study framework. I condition on time and state fixed effects and normalize the 2005 effect to zero. Because the health care reform was enacted July 2006, years are defined as July t to June $t+1$. For example, 2006 represents July 2006 to June 2007.

The results suggests a sharp downward pre-trend in Massachusetts relative to the national average. Overall, it would be difficult to argue that the parallel trends assumption holds in this case and that difference-in-differences can provide a useful estimate. Synthetic control estimation represents a possible solution to the differential trends implied by Figure 1.

3 Model

I build on the notation used in Section 2.1. The modified estimator will require synthetic controls for all units. The vector of weights to construct a synthetic control for unit i is represented by

$$\boldsymbol{w}_i = (w_1^i, \ldots, w_{i-1}^i, w_{i+1}^i, \ldots, w_N^i),$$

where w_j^i represents the weight given unit j for the creation of the synthetic control for unit i. The weights are constrained, as before, to be non-negative and to sum to one. I define the set of possible weighting vectors to create the unit i synthetic control by

$$\mathcal{W}_i = \left\{ \boldsymbol{w}_i \;\middle|\; \sum_{j \neq i} w_j^i = 1, w_j^i \geq 0 \text{ for all } j \right\}. \tag{6}$$

As before, the specification is represented by equation (1). The advantage of this model is that it allows for correlation between treatment assignment and $\boldsymbol{\lambda}_t \boldsymbol{\mu}_i$, permitting rich non-linear and non-smooth trends and outcome shocks in the treated unit.

In this section, I discuss the two modifications of this paper. First, I relax the convex hull assumption required by the SCM. Next, I modify the approach to permit transitory shocks to the outcome variable, represented by ϵ_{it} in equation (1). Note that we usually assume such transitory shocks; an assumption that the outcomes perfectly represent the underlying interactive fixed effects likely never holds in practice. The following sections consider these restrictions independently, followed by a more formal derivation of the full ISCM in Section 4.

3.1 Outside the Convex Hull

The synthetic control method assumes that the pre-treated outcomes of the treated unit are within the "convex hull" of the outcomes of the other units such that

$$\sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j = \boldsymbol{\mu}_1. \tag{7}$$

Researchers are encouraged to use the pre-period to test the appropriateness of the assumption in equation (7): "We do not recommend using this method when the pretreatment fit is poor..." (Abadie et al. (2015)). However, it is still possible to use synthetic control estimation even when this assumption fails. The insight in this section is to consider the treated unit as a possible control for untreated units.

3.1.1 Estimating synthetic controls for all units

The modified approach of this paper estimates synthetic controls for all units. This modification offers the opportunity to use an alternative assumption when appropriate:

$$\sum_{j\neq i}^{N} w_j^i \boldsymbol{\mu}_j = \boldsymbol{\mu}_i \quad \text{for some } i \text{ with } w_1^i > 0.$$

This condition states that the "convex hull" assumption holds for unit i and the treated unit is part of the synthetic control for unit i. Assuming either $\sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j = \boldsymbol{\mu}_1$ or the above condition, then it is possible to estimate the policy effect. For example, assume that $\sum_{j\neq i}^{N} w_j^i \boldsymbol{\mu}_j = \boldsymbol{\mu}_i$ holds for unit $i \neq 1$, then for $t > T_0$,

$$\begin{aligned} Y_{it} - \sum_{j\neq i}^{N} w_j^i Y_{jt} = & \ \boldsymbol{\lambda}_t \left(\boldsymbol{\mu}_i - \sum_{j\neq i}^{N} w_j^i \boldsymbol{\mu}_j \right) + \left(\alpha_{it} D_{it} - \sum_{j\neq i}^{N} w_j^i \alpha_{jt} D_{jt} \right) + \left(\epsilon_{it} - \sum_{j\neq i}^{N} w_j^i \epsilon_{jt} \right), \\ = & \ -\alpha_{1t} w_1^i + \left(\epsilon_{it} - \sum_{j\neq i}^{N} w_j^i \epsilon_{jt} \right). \end{aligned}$$

The first term on the right hand side, $\boldsymbol{\lambda}_t \left(\boldsymbol{\mu}_i - \sum_{j \neq i}^{N} w_j^i \boldsymbol{\mu}_j\right)$, is equal to zero by assumption. Since $D_{it} = 0$ for all $i > 1$, the second term reduces to $-\alpha_{1t} w_1^i D_{1t} = -\alpha_{1t} w_1^i$. Assuming that $w_1^i > 0$ and that D_{1t} is independent of ϵ_{it} for all i, then α_{1t} is identified by the above equation. Thus, even in cases in which there is no appropriate synthetic control for unit 1, unit 1 may receive a positive weight as part of the synthetic control for another unit. Then, it is possible to assign the change in the outcome to (the negative of) the fraction of the weight assigned to the treated unit.

The proposed estimation procedure creates synthetic controls for all units and regresses $Y_{it} - \sum\limits_{j \neq i} \hat{w}_j^i Y_{jt}$ on $D_{it} - \sum\limits_{j \neq i} \hat{w}_j^i D_{jt}$. The researcher can use the pre-period synthetic control fit to determine which units are appropriately within the convex hull of the other units and include these units in this regression. In this paper, I estimate the pre-period variance of the outcome relative to the synthetic control outcome:

$$\hat{V}_i \equiv \frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{it} - \sum_{j \neq i}^{N} \hat{w}_j^i Y_{jt} \right)^2 . \tag{8}$$

The final regression includes all units and is then weighted by $\frac{1}{\hat{V}_i}$. This approach downweights units with poor synthetic controls and places more weight on units with better fits.[5] However, alternative weighting procedures are also possible, including the use of a threshold for $\hat{V}_i$ to determine which units to use in the regression. In simulations below, I find that including all units and weighting by $\frac{1}{\hat{V}_i}$ performs well. Thus, the method of this paper is to construct synthetic controls for all units and then estimate the relationship between the difference in outcomes and difference in the policy variable using weighted least squares (WLS):

$$\widehat{\alpha}_{1t} = \underset{\alpha}{\operatorname{argmin}} \frac{1}{2N} \sum_{i=1}^{N} \frac{1}{\hat{V}_i} \left[\left(Y_{it} - \sum_{j \neq i} \hat{w}_j^i Y_{jt} \right) - \alpha \left(D_{it} - \sum_{j \neq i} \hat{w}_j^i D_{jt} \right) \right]^2 . \tag{9}$$

$\widehat{\alpha}_{1t}$ can be estimated for all time periods, including pre-intervention periods (setting $D_{1t} = 1$ for all t). Using the pre-intervention and post-intervention estimates is equivalent to estimation of an event study in a more traditional difference-in-differences setup. The pre-

[5] In the next section, I discuss further modifications to account for transitory shocks. Without transitory shocks, the measure in equation (8) is simply a measure of how close the assumption in equation (7) is to holding in practice.

intervention estimates provide evidence of the overall "fit" of the synthetic controls and a visual test for whether the method has accounted for systematic pre-trends and levels. This test is similar to checking the fit between the treated unit and its synthetic control in the pre-period using the SCM, but extended to account for the use of more information in generating the estimates.

It is straightforward to alter equation (9) to estimate more aggregated effects, such as one effect for the full post-intervention period:

$$\widehat{\alpha}_1 = \underset{\alpha}{\operatorname{argmin}} \frac{1}{2N\left(T-T_0-1\right)} \sum_{t=T_0+1}^{T} \sum_{i=1}^{N} \frac{1}{\hat{V}_i} \left[\left(Y_{it} - \sum_{j \neq i} \hat{w}_j^i Y_{jt} \right) - \alpha \left(D_{it} - \sum_{j \neq i} \hat{w}_j^i D_{jt} \right) \right]^2 .$$

While I have motivated the estimation of synthetic controls for all units by possible violations of the SCM's convex hull assumption, note that there are benefits to the proposed modification even in cases where the convex hull assumption holds. The modified approach uses more information and should improve power. Moreover, in practice, it is often difficult to precisely determine whether the convex hull assumption holds. Using more information and upweighting units in which the convex hull assumption is more likely to hold should improve inference.

3.1.2 Simulation Results

To illustrate the usefulness of this approach, I provide simulation results. I generate an outcome variable which is a function of three interactive fixed effects:

$$Y_{it} = \sum_{k=1}^{3} \lambda_t^{(k)} \mu_i^{(k)} + \epsilon_{it}, \tag{10}$$

$$\lambda_t^{(1)} = \begin{cases} 0.2t & \text{if} \quad t < 40 \\ 8 & \text{if} \quad t \geq 40 \end{cases}$$

$$\lambda_t^{(2)} = 0.05t$$

$$\lambda_t^{(3)} = \begin{cases} 0.2 - 0.02t & \text{if} \quad t < 20 \\ 0 & \text{if} \quad t \geq 20 \end{cases}$$

I assume $\mu_i^{(k)} \sim U(0,1)$ for all k and for all $i > 1$. For $i = 1$, $\left(\mu_1^{(1)}, \mu_1^{(2)}, \mu_1^{(3)}\right) = (1.0, 0.5, 0.4)$. I generate data for $N = 30, T = 50$ and define the policy variable as equal to 0, except for unit 1 in the final time period:

$$D_{it} = \begin{cases} 0 & \text{if } i > 1 \text{ or } t < 50 \\ 1 & \text{if } i = 1 \text{ and } t = 50. \end{cases}$$

The policy effect in equation (10) is equal to 0. Note that $\mu_1^{(1)} > \mu_j^{(1)}$ for all $j \neq 1$ which should create problems for the traditional synthetic control approach. The interactive fixed effects generate complicated within-unit trends. For these simulations, I set $\epsilon_{it} = 0$ for all (i, t) to focus on the convex hull assumption as the next section discusses transitory shocks. The simulation results are provided in Table 1. A typical fixed effects ("difference-in-differences") specification produces biased estimates given the correlation between treatment assignment and the interactive fixed effects. I estimate a bias of 1.592. I also estimate the specification while including a 5-degree polynomial in t for each state. The bias decreases, but I still find a mean bias of 0.166. Even a flexible polynomial does not adequately fit the underlying behavior of the outcome variable. Next, I implement the traditional synthetic control method using only unit 1 and its synthetic control. I find a bias equal to 0.082. However, when I use all states and equation (9), the resulting estimate is close to the true value as the mean bias decreases to -0.006.

When the SCM is applied only to the treated state, the method produces biased estimates when that state is outside of the convex hull. However, it is straightforward to apply the SCM to all states, which permits the treated state to be part of the synthetic control of other states. This extension provides additional information when estimating the policy effect and reduces the bias.

Table 1: Simulations: Not in Convex Hull

	Mean Bias	Median Absolute Deviation	RMSE
Fixed Effects	1.592	1.593	1.603
with Trends	0.166	0.166	0.167
SCM, Treated Only	0.082	0.090	0.313
SCM, All States	-0.006	0.009	0.022

The "fixed effects" estimator refers to a model with additive state and time fixed effects. When state-specific trends are included, a 5 degree polynomial in time is included for each state. "All States" means that synthetic controls are estimated for each unit and the policy effect is estimated using equation (9). RMSE is root mean squared error.

There are also gains to using this modified approach even when the convex hull assumption holds for the treated unit. I provide additional simulation results in which this restriction is likely to hold. I replicate the same data generating process as above with the lone exception of setting $\left(\mu_1^{(1)}, \mu_1^{(2)}, \mu_1^{(3)}\right) = (0.8, 0.5, 0.4)$. The results are presented in Table 2. Given that there are no transitory shocks and the treated unit is (likely) within the convex hull, the SCM fares very well with estimated bias equal to -0.0004. The modified approach introduced in this section recommends using all states, and this approach also performs well with a bias of -0.0001. The Median Absolute Deviation and RMSE metrics provide more evidence about the precision of the estimates, and the results suggest reductions when all states are used. The Median Absolute Deviation decreases by 17% and the RMSE decreases by 43%. By using all states, the estimator uses more information and appears to increase precision.

Table 2: Simulations: Convex Hull Assumption Likely Holds for Treated Unit

	Mean Bias	Median Absolute Deviation	RMSE
Fixed Effects	0.9557	0.9560	0.9732
with Trends	0.1018	0.1019	0.1033
SCM, Treated Only	-0.0004	0.0063	0.0159
SCM, All States	-0.0001	0.0052	0.0091

The "fixed effects" estimator refers to a model with additive state and time fixed effects. When state-specific trends are included, a 5 degree polynomial in time is included for each state. "All States" means that synthetic controls are estimated for each unit and the policy effect is estimated using equation (9). RMSE is root mean squared error.

3.2 Estimating Consistent Synthetic Controls Given Transitory Shocks

Abadie et al. (2010) discusses the usefulness of weights such that $\boldsymbol{\mu}_1 = \sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j$ but recognizes that since $(\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_N)$ are unobserved, this relationship must be estimated using the pre-intervention outcomes. The outcomes, however, are noisy measures of the variables of interest. This noise prevents consistent estimation of the weights. The reason for this bias is similar to the bias associated with classical measurement error in an explanatory variable with OLS estimation. In this section, I only consider estimating a synthetic control for the treated state, unit 1.

3.2.1 Two-step process

Starting with the Abadie et al. (2010) assumption, it is straightforward to derive the relationship between the outcomes of the treated unit and the outcomes of the controls units in each period using equation (1) for $t \leq T_0$:

$$
\begin{aligned}
\boldsymbol{\mu}_1 =& \sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j, \\
\boldsymbol{\lambda}_t \boldsymbol{\mu}_1 =& \boldsymbol{\lambda}_t \sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j, \\
Y_{1t} - \epsilon_{1t} =& \sum_{j=2}^{N} w_j^1 \left(Y_{jt} - \epsilon_{jt} \right), \\
Y_{1t} =& \sum_{j=2}^{N} w_j^1 Y_{jt} + \left(\epsilon_{1t} - \sum_{j=2}^{N} w_j^1 \epsilon_{jt} \right).
\end{aligned}
$$

Without considering the constraints on the weights, note that a regression of Y_{1t} on $(Y_{2t}, \ldots, Y_{Nt})$ does not produce consistent estimates of the weights since Y_{jt} is correlated with the error term $\epsilon_{1t} - \sum_{j=2}^{N} w_j^1 \epsilon_{jt}$ (when $w_j^1 > 0$):

$$E\left[Y_{1t}\middle|Y_{2t},\ldots,Y_{Nt}\right] = E\left[\sum_{j=2}^{N} w_j^1 Y_{jt} + \left(\epsilon_{1t} - \sum_{j=2}^{N} w_j^1 \epsilon_{jt}\right)\middle|Y_{2t},\ldots,Y_{Nt}\right] \neq \sum_{j=2}^{N} w_j^1 Y_{jt},$$

due to the correlation between Y_{jt} and ϵ_{jt}. A formal argument of this bias under general conditions is provided in Ferman and Pinto (2016), which also discusses (restrictive) conditions under which the SCM will produce consistent weights.[6]

In this paper, I start with the same assumption as Abadie et al. (2010) but in an expected value framework. I assume that there exists weights w_j^1 such that

$$E\left[\boldsymbol{\mu}_1\right] = \sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j.$$

I also assume that the observation-specific error term is mean-zero conditional on time for all i:

$$E\left[\epsilon_{it}|t\right] = 0. \tag{11}$$

Next, define predicted values of Y as a function of unit-specific trends:

$$\widehat{Y}_{jt} \equiv \hat{f}_j(t). \tag{12}$$

For example, one possible choice of $f_j(t)$ is a polynomial such that

$$\widehat{Y}_{jt} \equiv \widehat{\gamma}_j + \sum_{k=1}^{K} \widehat{\beta}_{jk} t^k,$$

where only the pre-intervention period is used to estimate these trends. The idea behind the imperfect synthetic control method is that these smooth trends are often able to imperfectly predict the outcome variable trajectory. This approach is equivalent to an instrumental variable method where the state-specific time trends are used as "instruments" to predict

[6]Ferman and Pinto (2016) suggest an "IV-Like SC Estimator" which uses lagged values of the outcome variables as instruments though this approach requires "[i]mposing strong assumptions on the structure of the idiosyncratic error..." The approach suggested in this section does not impose such assumptions.

changes in the outcome variable. The "first stage" assumption is that these trends are predictive such that for $t \leq T_0$,

$$E\left[Y_{jt}|\widehat{Y}_{jt}\right] = \widehat{Y}_{jt}.$$

The assumption is *not* that $\hat{f}_j(t)$ perfectly predicts Y_{jt} for all time periods but only that it is correlated with Y_{jt}. Y_{jt} may follow a less smooth trajectory than required for a polynomial, but a flexible unit-specific function can predict functions correlated with non-smooth functions. Other functions are also possible and permitted under equations (11) and (12).[7] I also assume that $\widehat{Y}_{kt}$ does not provide additional information about Y_{jt} after conditioning on $\widehat{Y}_{jt}$, though this assumption is only a simplification. This simplification modifies the above condition slightly:

$$E\left[Y_{jt}|\widehat{Y}_{1t}, \ldots, \widehat{Y}_{Nt}\right] = \widehat{Y}_{jt} \quad \text{for } t \leq T_0. \tag{13}$$

Under the assumptions contained in equations (11) and (13),

$$E\left[Y_{1t}\middle|\widehat{Y}_{2t}, \ldots, \widehat{Y}_{Nt}\right] = E\left[\sum_{j=2}^{N} w_j^1 Y_{jt} + \left(\epsilon_{1t} - \sum_{j=2}^{N} w_j^1 \epsilon_{jt}\right)\middle|\widehat{Y}_{2t}, \ldots, \widehat{Y}_{Nt}\right] = \sum_{j=2}^{N} w_j^1 \widehat{Y}_{jt}. \tag{14}$$

This derivation suggests that it is possible to estimate consistent weights using the predicted values of the outcome. The weights are estimated using

$$\hat{\boldsymbol{w}}_1 = \underset{\boldsymbol{\phi}_1 \in \mathcal{W}_1}{\operatorname{argmin}} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j \neq i} \phi_j^1 \widehat{Y}_{jt}\right)^2 \right\}, \tag{15}$$

which imposes that the weights belong to the set as defined in equation (6). Importantly, condition (13) does not presume that it holds for $t > T_0$. Such an assumption would require that the $\hat{f}_j(t)$ function accurately predicts the treated unit's post-intervention counterfactual outcomes. Instead, condition (13) simply assumes that the $\hat{f}_j(t)$ function provides useful information on average about each unit's pre-intervention outcomes. The predicted function helps trace out the outcome trajectory for the unit, even when this trajectory is not smooth, and is useful as long as that smooth approximation is correlated over the pre-period with the

[7]For example, given longer pre-periods, it may make sense to use splines.

outcome variable. Given the weights estimated using equation (15), then the policy effect in time period t can be estimated using equation (4).

3.2.2 Modification

It is possible to replace equation (15) with

$$\hat{\boldsymbol{w}}_1 = \underset{\boldsymbol{\phi}_1 \in \mathcal{W}_1}{\operatorname{argmin}} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} \left(\widehat{Y}_{1t} - \sum_{j \neq i} \phi_j^1 \widehat{Y}_{jt} \right)^2 \right\}, \tag{16}$$

which minimizes the pre-period mean squared error between predicted values of the outcome variable. The advantage of this modification is that it may be more appropriate to use the predicted value for both unit i and the other units when constructing the synthetic control weights. The use of the predicted values is appropriate given equations (13) and (14).

The value of this approach is that $\hat{f}_j(t)$ only serves to trace out the outcome trajectory of unit j to an approximation and may do so poorly in some time periods and better in other time periods. To the extent that there are large discontinuities in the pre-period, polynomials will smooth out these shocks. A synthetic control approach is valuable in such a case because it potentially could estimate a synthetic control which experienced similar pre-intervention discontinuities. However, it may be less desirable to estimate this synthetic control if predicted values, generated using a smooth function in t, are used for the untreated units, but the treated unit's outcomes are not smoothed in the same manner. If the treated unit's outcomes are also replaced by a similar smooth function, then the predicted outcomes between the treated unit and synthetic unit should be more comparable. In general, both equation (15) and (16) produce appropriate synthetic controls; the simulations below will use equation (16).

3.2.3 Simulation Results

I test the usefulness of the two-step procedure to the synthetic control approach by presenting simulation results. I use the same data generating process as introduced in Section 3.1.2 with two modifications. First, I set $\left(\mu_1^{(1)}, \mu_1^{(2)}, \mu_1^{(3)}\right) = (0.8, 0.5, 0.4)$ such that the convex hull assumption is plausible. Second, there is now an error term with positive variance. I use $\epsilon_{it} \sim N(0, 1)$. The results are presented in Table 3. As before, the fixed effects regression produces biased estimates due to a correlation between the unobserved interactive fixed effects and treatment assignment. Including state-specific trends reduces this bias to 0.090.

The traditional synthetic control method has mean bias of 0.237. I implement the imperfect synthetic control approach of this paper by first predicting the outcomes for each unit using a 5-degree polynomial in t. These predictions are then used to estimate the synthetic control weights. This two-step process reduces the bias to 0.000.[8]

Table 3: Simulations: With Error

	Mean Bias	Median Absolute Deviation	RMSE
Fixed Effects	0.951	1.019	1.408
with Trends	0.091	0.976	1.454
SCM, Treated Only	0.237	0.797	1.186
ISCM, Treated Only	0.000	0.791	1.197

The "fixed effects" estimator refers to a model with additive state and time fixed effects. When state-specific trends are included, a 5 degree polynomial for each state is included. RMSE is root mean squared error.

4 Imperfect Synthetic Control Method

In this section, I introduce the "imperfect synthetic control method" (ISCM), which combines the modifications presented in Sections 3.1 and 3.2. It uses all units in the manner discussed in Section 3.1, reducing concerns that the method cannot create an appropriate control for a unit without an appropriate synthetic control. It also creates the synthetic control for each unit using predicted values of the outcome variable, instead of the actual values.

4.1 Conditions

ISCM relies on the following assumptions:

[8]Note that even though ISCM has lower bias than the traditional SCM, the SCM has lower variance (as measured by RMSE). This result is not necessarily surprising given that the SCM is designed to minimize the mean squared error in the pre-period (given the constraints on the weights), which *may* have ramifications for the mean squared error in the post-period.

A1 (Outcomes): $Y_{it} = \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \alpha_{it} D_{it} + \epsilon_{it}$, where D_{it} is specified in equation (2).

A2 (Existence of Synthetic Controls): (a) There exists $\boldsymbol{w}_1 \in \mathcal{W}_1$ such that $E[\boldsymbol{\mu}_1] = \sum_{k \neq 1} w_k^1 \boldsymbol{\mu}_k$.
or (b) There exists $\boldsymbol{w}_j \in \mathcal{W}_j$ with $w_1^j > 0$ such that $E[\boldsymbol{\mu}_j] = \sum_{k \neq j} w_k^j \boldsymbol{\mu}_k$.

A3 (Independence): $E\left[\epsilon_{it} | t\right] = 0$ and $E\left[\epsilon_{it} | D_{it}\right] = 0$.

A4 (First Stage): $E\left[Y_{jt} | \widehat{Y}_{1t}, \ldots, \widehat{Y}_{Nt}\right] = \widehat{Y}_{jt}$ for all j, where $\widehat{Y}_{jt}$ is defined in equation (12).

A5 (Within-Unit Dependence): (a) For each i, Y_{it} is a strongly mixing sequence in t with α of size $-\frac{r}{r-2}$, $r > 2$; (b) $E|Y_{it}|^{r+\delta}$ for some $\delta > 0$.

A6 (Identification of weights):

$$\begin{bmatrix} \widehat{Y}_{11} & \ldots & \widehat{Y}_{N1} \\ \vdots & \ddots & \vdots \\ \widehat{Y}_{1T_0} & \ldots & \widehat{Y}_{NT_0} \end{bmatrix} \text{ is rank } N.$$

Assumptions **A1-A4** generalize those presented in the previous sections. **A5** permits within-unit dependence and non-stationarity. **A6** is a straightforward assumption to ensure that the synthetic control weights are unique (and assumes $T_0 \geq N$). If **A6** does not hold, then it may be possible to use a subset of units in which the condition holds. Alternative assumptions are also possible.[9] Similar to before, for each i,

$$\hat{\boldsymbol{w}}_i = \underset{\boldsymbol{\phi}_i \in \mathcal{W}_i}{\operatorname{argmin}} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} \left(\widehat{Y}_{it} - \sum_{j \neq i} \phi_j^i \widehat{Y}_{jt} \right)^2 \right\} \tag{17}$$

The policy effect at time t is estimated as

$$\hat{\alpha}_{1t} = \underset{b}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\hat{V}_i} \left(Y_{it} - bD_{it} - \sum_{j \neq i} \hat{w}_j^i \left(Y_{jt} - bD_{jt} \right) \right)^2, \tag{18}$$

given $\hat{V}_i$ defined in equation (8). The above equation simply represents a weighted least squares regression. The ISCM can be summarized in the following steps:

[9]It is not clear that unique synthetic control weights are necessary for proper estimation of the policy effect.

1. For each i, regress Y_{it} on a function of t and a constant in the pre-treatment period.
2. Set $\widehat{Y}_{it} \equiv \hat{f}_i(t)$ for all i.
3. Estimate weights $\boldsymbol{w}_j$ for all j using equation (17). This step can be implemented with the `synth` package in Stata.
4. Calculate the parameter estimate using equation (18).

Overall, the ISCM is simple to implement, requiring only WLS estimation and use of the `synth` package.

4.2 Simulation Results

I present simulation results using similar data as before. These simulations generate a treated state which is not in the convex hull of the other states. This property is implemented by using $\left(\mu_1^{(1)}, \mu_1^{(2)}, \mu_1^{(3)}\right) = (1.0, 0.5, 0.4)$. I also include an error, $\epsilon_{it} \sim N(0,1)$. The results are presented in Table 4. The fixed effects specification produces estimates with mean bias of 1.587. Including state-specific trends reduces the mean bias to 0.155. Next, I implement the traditional SCM, using only the treated unit and its synthetic control. The mean bias is 0.738. Using all states, as suggested in Section 3.1, decreases this bias to 0.142.

Relative to the traditional SCM, using the two-step approach suggested in Section 3.2 reduces the mean bias from 0.738 to 0.248. Finally, I implement the suggested method of this paper, using all units and the two-step process. The mean bias for this approach is -0.001, a substantial improvement on the other methods. For these simulations, ISCM also performs better than SCM in terms of median absolute deviation and RMSE.

4.3 Properties

There is only one treated unit so $\widehat{\alpha}_{1t}$ does not converge to its true value. Instead, I show that, under the given conditions, the expectation of the estimate converges to its true value. This property is formalized below and in the Appendix.

Theorem 4.1 (Asymptotically Unbiased)**.** *If* ***A1-A6*** *hold, then* $E\left[\widehat{\alpha}_{1t} - \alpha_{1t}\right] \xrightarrow{p} 0$ *as* $T_0 \to \infty$.

Table 4: Simulations: Not in Convex Hull, With Error

	Mean Bias	Median Absolute Deviation	RMSE
Fixed Effects	1.587	1.580	1.897
with Trends	0.155	0.976	1.459
SCM, Treated Only	0.738	0.971	1.419
SCM, All States	0.142	0.904	1.360
ISCM, Treated Only	0.248	0.879	1.329
ISCM, All States	-0.001	0.889	1.322

The "fixed effects" estimator refers to a model with additive state and time fixed effects. When state-specific trends are included, a 5 degree polynomial for each state is included. RMSE is root mean squared error.

Note that it is common – including in Section 5 below – to estimate an aggregate treatment effect instead of one per post-treatment time period. Given a long post-period ($T - T_0 + 1$), the estimate $\widehat{\alpha}$ (as defined in equation (9)) converges to the true average of the period-specific estimates.

4.4 Measuring Fit

It is typical when using the SCM to examine the fit of the synthetic control in the pre-intervention period. An equivalent test can be implemented with the proposed method by estimating α_{1t} for each $t \leq T_0$ using equation (18). This approach is similar to an event study approach which estimates the "effects" before the intervention to test for pre-existing trends. In this case, the approach also tests for pre-existing level differences since the ISCM should also reduce systematic pre-intervention level differences.

4.5 Inference

4.5.1 Across Unit Permutation Tests

For inference, Abadie and Gardeazabal (2003) and Abadie et al. (2010) recommend a permutation test in which a treatment effect is estimated for each unit in the donor pool as if it were treated (and unit 1 is included in the donor pool). Under the null hypothesis that there is no treatment effect, the idea is that these placebo tests generate a distribution of estimates which would be observed randomly. The placement of $\hat{\alpha}$ in this distribution generates an appropriate p-value under the assumption that the error term is identically distributed across units, similar to the assumption required for the method proposed in Conley and Taber (2011).[10] Ando and Sävje (2013) discusses possible limitations of the placebo test approach for synthetic control estimation. A major concern with the permutation test approach is that while it is typical to test whether an appropriate synthetic control exists for the treated unit (by studying the pre-intervention fit), this verification is not often done for each placebo estimate. Thus, some units may not have a valid synthetic control and the placebo estimates themselves are not accurate realizations of the underlying estimate distribution.

The modified approach introduced in this paper should reduce these concerns. The motivation of ISCM is to relax restrictions required by the synthetic control approach. However, a complementary benefit is that it should produce more valid placebo estimates as well. There are still three limitations of this inference approach. First, as mentioned, it assumes that the error term is identically-distributed across units. This assumption is especially restrictive given that different units, such as states or countries, often have different population sizes so the outcomes should have different variances. Ferman and Pinto (2015) discusses adjustments (in the difference-in-differences and synthetic control cases) when the source of heteroscedasticity (such as population size) can be modeled.

Second, while ISCM reduces concerns about units not having appropriate synthetic controls, it does not eliminate them, suggesting that the placebo estimates may not represent the true underlying distribution of the estimates under the null hypothesis. Finally, the placebo estimates themselves are not independent. For example, consider a case in which unit i's synthetic control is unit j (i.e., $w_j^i = 1$) and unit j's synthetic control is unit i (i.e., $w_i^j = 1$). Then, the placebo estimates for period T include $\epsilon_{iT} - \epsilon_{jT}$ and $\epsilon_{jT} - \epsilon_{iT}$, which are equal in magnitude (it is often suggested to compare the absolute value of the estimated effect to the absolute value of the placebo estimates) and clearly dependent. Generally, a synthetic control is composed of multiple units which means that there is dependence across several, and likely all, of the placebo estimates. This is problematic since each placebo estimate does not represent independent information about the distribution of the estimated policy effect.

[10] Alternate approaches and test statistics have also been mentioned in the literature (see Firpo and Possebom (2016) and Hahn and Shi (2017) for further discussions). These approaches likely also apply to ISCM with similar concerns as discussed in the above text.

4.5.2 Within-Unit Permutation Tests

Alternatively, Chernozhukov et al. (2018) recommends only using the treated unit's residuals (defined as the difference in the treated unit's outcomes and the predicted outcomes for each time period), using the pre-treatment periods to generate placebo estimates. This approach can also be used with ISCM. The benefit of this approach in the synthetic control context is that it does not suffer from any of the three limitations mentioned above since it only uses the treated unit (and its synthetic control). Instead, this approach requires a stationarity assumption on the errors, requiring the distribution of the error term to be independent of time.

Moreover, while the approach can be implemented to test sharp null hypotheses, it may not be useful to test hypotheses for the average effect, which is often of interest. To generate counterfactual average effects, the approach creates estimates using time periods before treatment that are the same length as the post-treatment period. However, when the post-treatment period is large or, more accurately, not small relative to the pre-period, there are limited numbers of counterfactual average effects. For example, in the application of this paper, the post-period includes 56 months while the pre-period is 124 months, implying that there are only 2 counterfactual average effects that can be generated using the pre-period, which is inadequate for the calculation of an informative p-value.

4.5.3 Inference when the post-treatment period is large

When the post-period is large, it is possible to apply inference methods using only the post-treatment data for the treated unit. These methods maintain the benefits of the Chernozhukov et al. (2018) approach. However, in addition, these approaches will not require a stationarity assumption. Given a long post-period, it is possible to only use the post-treatment data of the treated unit, avoiding restrictions on the distribution of the errors across units or over time. The idea is to consider the period-by-period estimates as a time series. For example, one can apply Ibragimov and Müller (2010) or Canay et al. (2017) to the estimated post-treatment effects. Ibragimov and Müller (2010), building on results in Bakirov and Székely (2006), introduces a method which partitions the data into $q \geq 2$ groups and estimates the policy effect separately in each group. Using the q estimates, it is possible to construct a t-statistic and obtain an estimated p-value. This approach is conservative when there is heteroscedasticity and only valid for $q \leq 14$ at significance level 5%. In one set of simulations, Ibragimov and Müller (2010) partitions time series data into 16 sets of 8 consecutive observations. Longer partitions permit more serial correlation.

Canay et al. (2017) adopts a similar approach, partitioning the data and creating a test statistic from the q estimates. Instead of comparing the resulting t-statistic to standard critical values, they simulate the distribution of the test statistic using Rademacher weights. This approach is valid for all $q \geq 2$ and significance levels, and it is not conservative. In this paper, I use the non-randomized version of the Canay et al. (2017) approach (see that paper for details), which can be conservative when q is small. In one set of simulations, Canay et al. (2017) partition time series data into 12 groups of 8 (or 9) consecutive observations and find that the inference procedure works well even with high levels of serial correlation. In the application below, I adopt this inference procedure.

The test statistic for hypotheses governing the average effect, given the estimated effect $\widehat{\beta} = \frac{1}{q}\sum_{j=1}^{q}\widehat{\beta}_j$, is

$$T = \frac{|\widehat{\beta} - \beta_0|}{s_{\hat{\beta}}/\sqrt{q}},$$

where β_0 is the policy effect under the null hypothesis and $s_{\hat{\beta}} = \sqrt{\frac{1}{q-1}\sum_{j=1}^{q}\left(\widehat{\beta}_j - \widehat{\beta}\right)^2}$ given $\widehat{\beta}_j$ representing the estimate for partition j. Simulating the distribution of this test statistic simply repeats the creation of this test statistic using $W_j\left(\widehat{\beta}_j - \beta_0\right)$ where W_j represents an independent draw from the Rademacher distribution. The placement of T in the distribution of simulated test statistics provides the p-value. When the number of partitions is small, it may be more appropriate to use the weights introduced in Webb (2013).

4.6 Extensions

4.6.1 Multiple Treated Units

Following the development of the SCM, I have assumed that there is only one treated unit. However, it is common to study policies adopted by many units. It is straightforward to extend ISCM to these empirical applications and, in fact, potentially simpler when compared to the SCM since ISCM already requires creating synthetic controls for every unit. If we assume that multiple units adopted the policy of interest at the same time, then implementation of ISCM is nearly-identical to the case in which only one unit adopts the policy.

Estimation of the policy effect $\hat{\alpha}_t$ can use equation (18) where D_{it} may equal 1 for multiple units when $t > T_0$. To permit state-specific heterogeneity in the policy effect (i.e, estimating α_{it} instead of α_t), then it is straightforward to modify this objective function. Without loss of generality, assume that the first N_1 units are treated. Then,

$$(\hat{\alpha}_{1t}, \cdots, \hat{\alpha}_{N_1,t}) = \underset{b_1,\cdots,b_{N_1}}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\hat{V}_i} \left(Y_{it} - b_i D_{it} - \sum_{j \neq i} \hat{w}_j^i \left(Y_{jt} - b_j D_{jt} \right) \right)^2. \tag{19}$$

The same approach can be used given differential timing of policy adoption but with a critical caveat. When estimating the synthetic control weights for unit i, it is important that the pre-intervention period cannot contain any treated observations. Equation (16) assumes that all possible control units are untreated. One possibility is to use the time period that is untreated for all units as a common pre-intervention period. An alternative is to exclude early-adopters as part of the donor pool for late adopters of the policy (i.e., assuming a weight equal to zero for the early-adopters), permitting the pre-intervention period to vary across units.

4.6.2 Adding Covariates

The specification in Abadie and Gardeazabal (2003) and Abadie et al. (2010) includes additional covariates (Z_{it}) such that $Y_{it} = \alpha_{it} D_{it} + Z_{it}' \Pi + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \epsilon_{it}$. It is assumed that synthetic control weights exist such that $\sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j = \boldsymbol{\mu}_1$ (as in equation (7)) *and* $\sum_{j=2}^{N} w_j^1 Z_{jt}' = Z_{1t}'$. However, Kaul et al. (2015) shows that the traditional SCM will not use the covariates to estimate the synthetic control weights if each period's outcome is used to generate the weights. Modifying equation (16) slightly permits estimation of weights which also account for variation in the covariates. Assume that there are L covariates, where I drop the time subscript so that the covariates are themselves defined by the time period (e.g., unemployment rate in period s). If one uses all pre-intervention outcomes to generate weights, then also using pre-intervention values of the covariates in the suggested manner will actually hurt the synthetic control fit. However, it may be beneficial to include post-treatment values of the covariates so this notation avoids committing to including covariates for specific time periods. Then,

$$\hat{\boldsymbol{w}}_i = \underset{\boldsymbol{\phi}_i \in \mathcal{W}_i}{\operatorname{argmin}} \left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} \left(\widehat{Y}_{it} - \sum_{j \neq i} \phi_j^i \widehat{Y}_{jt} \right)^2 + \frac{1}{L} \sum_{\ell=1}^{L} \left(Z_i^\ell - \sum_{j \neq i} \phi_j^i Z_j^\ell \right)^2 \right\}.$$

Note that it is also necessary to modify assumption **A2** to add that there exist weights such that the current **A2** holds as well as $\sum_{j=2}^{N} w_j^i Z_j' = Z_i'$. This approach "differences out" the covariates by imposing the assumption that the synthetic control also has the same covariates. If the Π coefficients are of interest (or if one does not want to impose this additional assumption), then it is possible to use Powell (2016) to jointly estimate the treatment effects and the parameters associated with the covariates.

5 Massachusetts Health Care Reform

Massachusetts passed comprehensive health care reform in 2006. The reform included a major Medicaid expansion, subsidies for low-income households, and implementation of an individual mandate. The reform served as a model for Affordable Care Act (ACA) and there is substantial interest in the effects of the Massachusetts reform given its similarities to the ACA.

In this section, I use geocoded data from the National Vital Statistics System (NVSS), a census of death in the United States. I construct mortality rates by state and month for ages 0-64 from January 1999 to February 2011. I study mortality amenable to health care using the ICD-10 codes in Sommers et al. (2014), though I will also show results for all mortality. I start the sample in January 1999 because ICD-9 codes were used before 1999. I end the sample in February 2011 so that the inference method has 7 equal-sized post-treatment partitions of 8 months. The results are insensitive to using earlier or later time periods, which I will show. The outcome variable is the health care amenable mortality rate (per 100,000) for ages 0 to 64. The traditional difference-in-differences specification is:

$$M_{it} = \alpha_i + \gamma_t + \beta \left[\mathbf{1}(i = \text{Massachusetts}) \times \mathbf{1}(t \geq \text{July 2006}) \right] + \epsilon_{it},$$

where M_{it} represents the mortality rate for state i in month t. July 2006 is the enactment date. For inference, I adapt the Canay et al. (2017) method, as discussed in Section 4.5. I set

$q = 7$ where each group includes 8 months (the post-period is July 2006 to February 2011).[11] I use this approach for a large set of null hypotheses and then invert the test statistic to generate 95% confidence intervals. It is important to note that this inference procedure does not necessarily produce symmetric confidence intervals.[12] To predict outcomes for ISCM, I fit a restricted cubic spline with 5 knots (Harrell (2001)) for each state.

[11] Regressions of the mortality rate on the lagged mortality rate in Massachusetts in the pre-period, and separately, in the post-period both generate estimates smaller than 0.1.

[12] I also tested the sharp null hypothesis that there was no mortality impact in any post-period month using the method proposed in Chernozhukov et al. (2018). I can reject the null hypothesis.

Table 5: Mortality Effects of Massachusetts Health Care Reform

Panel A: Mortality Ages 0-64 Per 100,000

				ISCM	
	Fixed Effects	SCM	SCM (All States)	Treated Only	All States
Estimate	-0.465	-0.779	-0.874	-0.621	-0.609
	[-0.675, -0.270]	[-1.082, -0.507]	[-1.233, -0.560]	[-0.959, -0.252]	[-0.950, -0.275]

Panel B: Alternate Outcomes

	ISCM, All States		
	Log(Mortality)	All Mortality	"Other Cause" Mortality
Estimate	-0.057	-0.546	-0.003
	[-0.082, -0.034]	[-0.862, -0.207]	[-0.274, 0.226]

Panel C: Alternate Time Periods and Inference

	ISCM, All States			
	Webb Weights	Longer Post-Period	Larger Partitions	Ibragimov and Müller (2010)
Estimate	-0.609	-0.664	-0.639	-0.639
	[-0.933, -0.322]	[-0.979, -0.355]	[-1.099, -0.331]	[-0.927, -0.352]

Notes: In Panel A, the column header refers to the approach used to generate the estimate. The outcome is health care amenable mortality per 100,000. 95% confidence intervals are presented in brackets, generated using the Canay et al. (2017) discussed in the text.

In Panel B, all estimates are generated using ISCM with all states. 95% confidence intervals are presented in brackets, generated using the Canay et al. (2017) discussed in the text. In the first column, I use log(health care amenable mortality rate) as the outcome variable. In the second column, I use all deaths per 100,000 as the outcome. In the third column, as a placebo test, I use "other cause" mortality, deaths that should not be affected by health care access.

In Panel C, the outcome is always health care amenable mortality per 100,000 and all estimates are generated using ISCM with all states. 95% confidence intervals are presented in brackets. In the first column, I use the weights recommended in Webb (2013) when the number of clusters to simulate the test statistic. In column 2, I extend the post-period to October 2011, which adds another partition for inference. In column 3, I lengthen the size of each partition to 12 months. The post-period extends to June 2011 so that there are 5 equal-sized partitions. In column 4, I length the size of each partition further to 20 months. Because there are only 3 partitions, I use the approach introduced in Ibragimov and Müller (2010).

The results are presented in Table 5, Panel A. The difference-in-differences estimate is -0.465, implying 0.465 fewer (health care amenable) deaths per 100,000 people. The event study equivalent was previously shown in Figure 1 and suggests that this estimate is driven primarily by pre-existing trends when Massachusetts is compared to the national average.

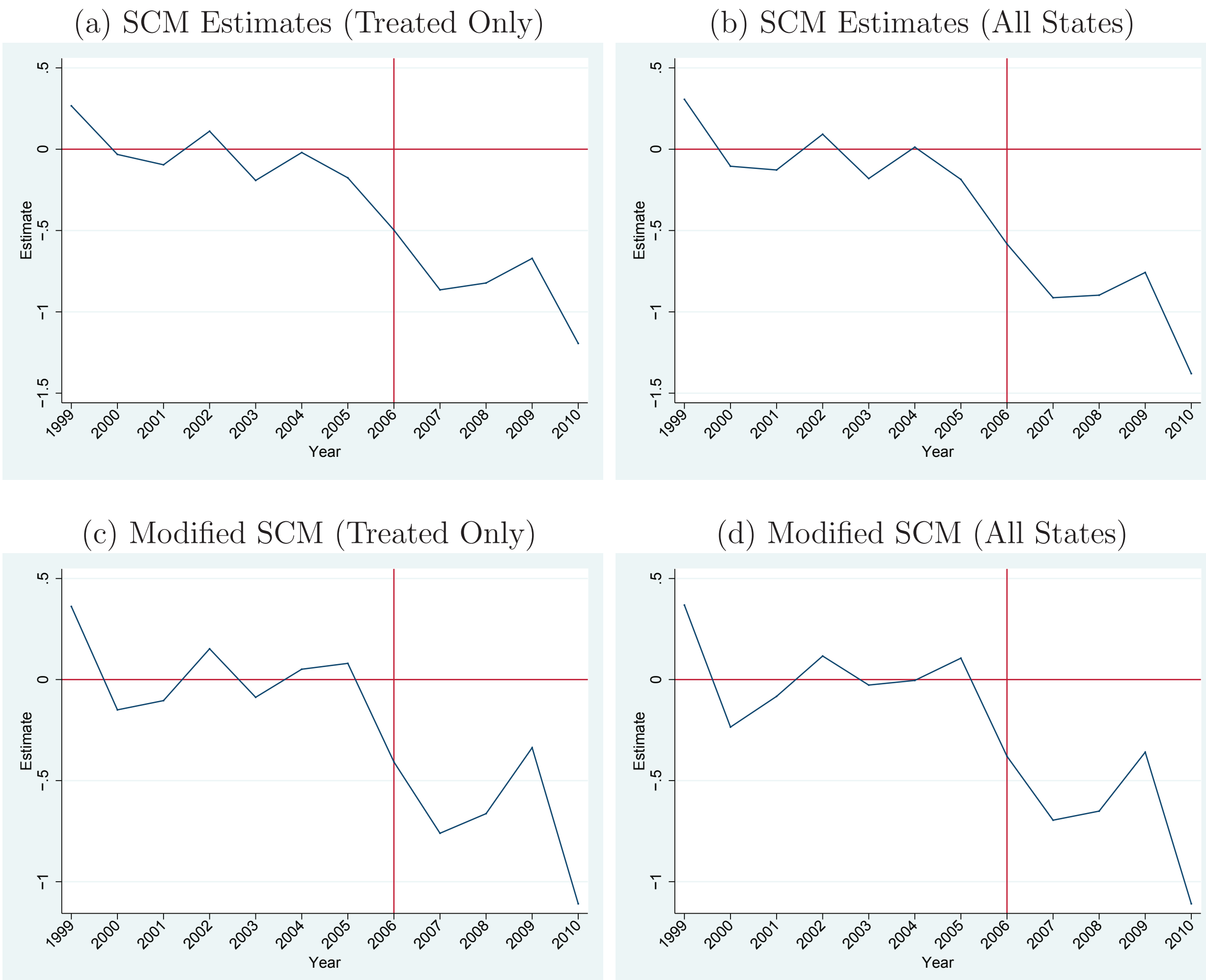
In the next column, I use the traditional SCM and estimate an effect of -0.779. I can reject that there is no mortality effect at the 5% significance level. The next column uses the traditional SCM but uses all states, corresponding to the approach discussed in Section 3.1. The estimate is similar. The event studies for these estimates are provided in the top half of Figure 2. Because state-level monthly mortality rates are volatile, I estimate effects at the annual level in all event studies. I define "year t" as July t to June $t+1$ for the purposes of these graphs since the Massachusetts reforms began July 2006. There is still some (though less) evidence of a slight downward trend before the reforms in the SCM event studies.

The final two columns use the ISCM. First, I generate the policy effect estimate using only Massachusetts and its synthetic control. The estimate is -0.621, statistically significant from zero at the 5% level. The final column presents the preferred estimate of the paper. I estimate that the Massachusetts health care reform reduced the mortality rate by 0.609 deaths per 100,000. The predicted (counterfactual) mortality rate for Massachusetts after the reforms was 11.48 deaths per 100,000, implying that the reforms reduced the (health care amenable) mortality rate by 5.3%. The ISCM event study estimates are presented in the bottom half of Figure 2. In both graphs, there is little evidence of any pre-existing trends and the average pre-reform estimates are close to zero (e.g., the corresponding pre-reform estimate related to Figure 2(d) is only 0.016).

Using the same weights and regression approach to generate the ISCM (all states) estimate, I study the implied differences in health insurance coverage over time using annual data from the Current Population Study. I present the event study in Appendix Figure A.1. I observe a relative increase in health insurance coverage similar to increases estimated in the literature (e.g., Miller (2012); Sommers et al. (2014)).

In Panel B, I include estimates for alternative outcomes. In the first column, the outcome is the log of (health care amenable) mortality. I estimate the Massachusetts health care reform reduced the mortality rate by 5.7%, similar to the implied percentage estimate when estimated in levels. In the second column, I replicated the main estimate of this paper but for all mortality, not just the health care amenable conditions defined in Sommers et al.

Figure 2: Synthetic Control Event Study Estimates



Notes: I estimate the policy effect for each year, both before and after enactment in an event study framework. Because the health care reform was enacted July 2006, years are defined as July t to June $t+1$. For example, 2006 represents July 2006 to June 2007.

(2014). The estimate is similar and implies a 3% reduction in the mortality rate given the higher baseline when all deaths are included.[13] In the final column of Panel B, I study "other cause" deaths, those not considered amenable to health care. This outcome acts as a valuable placebo test. I do not find any evidence of mortality reductions for this category, suggesting that the main estimates are not simply picking up some confounding Massachusetts-specific shock that reduced mortality.

In Panel C, I study the effects of varying the length of the time period while also testing the robustness of confidence intervals to changes to the inference method. Because I only have 7 partitions, it may be more appropriate to use the Webb (2013) weights to perturb the test statistic and generate p-values. The first column of Panel C applies these weights in the inference procedure while repeating the main estimate. The confidence intervals are similar. In the next column, I lengthen the post-period to October 2011, which permits an additional partition for inference. The estimate is slightly larger in magnitude and the 95% confidence interval excludes any values implying less than a reduction of 0.355 deaths per 100,000. In the next column, I define the post-period as July 2006 to June 2011 and partition the data into 5 groups of size 12 months. The larger-sized groups permit more serial correlation using the Canay et al. (2017) method. The estimated effect size is similar and is statistically significant from zero at the 5% level. In the final column, I use the same time period (i.e., the same point estimate) but use 3 partitions (each partition is 20 months). With only 3 partitions, it is more appropriate to use the (conservative) approach suggested in Ibragimov and Müller (2010). The confidence intervals are similar to the main confidence intervals, suggesting that they are not driven by inappropriately assuming less serial correlation than exists in the data.

6 Conclusion

Many empirical applications rely on panel data to compare the outcomes of treated units to the outcomes of untreated units. It is common to condition on unit and time fixed effects, implicitly assuming that the average unit is an appropriate control for each treated unit. However, the synthetic control method of Abadie and Gardeazabal (2003) and Abadie et al.

[13]The method generates entirely new synthetic control weights when the outcome changes. Alternatively, I could use the same weights generated for health care amenable mortality and applied to all mortality. The estimate is similar.

(2010, 2015) relaxes this assumption, permitting state-level trends and outcome shocks to be correlated with treatment assignment. While the SCM has proven quite useful in applied work, it requires two major assumptions. These restrictions are relaxed in this paper.

The ISCM of this paper, first, relaxes the convex hull assumption of the SCM. It relies on the insight that the treated unit may be part of the synthetic control of one or more control units, and the post-adoption outcome differences in these units also provide estimates of the policy effects. The ISCM develops synthetic controls for all units and then aggregates this information to estimate the policy effect. Even when the convex hull assumption holds for the treated unit, there are gains in using more states.

Second, it has been shown that the synthetic control weights are biased given that they are estimated to minimize the distance between outcomes in the pre-period and these outcomes are noisy measures of the underlying trends and shocks of interest. In this paper, I propose a two-step process in which the outcomes are estimated using simple state-specific trend terms. These predicted values are then used to estimate the synthetic control weights. These weights are consistent under straightforward assumptions.

The proposed estimator works well in simulations with less bias than the traditional SCM. I provide evidence of the usefulness of the estimator by estimating the effect of the Massachusetts health care reform on mortality rates. I estimate that the reform decreased mortality by about 3%.

Appendix

Theorem 4.1 (Asymptotically Unbiased)**.** *If **A1**-**A6** hold, then* $E\left[\widehat{\boldsymbol{\alpha}}_{1t} - \boldsymbol{\alpha}_{1t}\right] \xrightarrow{p} 0$ *as* $T_0 \to \infty$.

Proof. The parameter set is compact by assumption. $\left(Y_{it} - \sum\limits_{j \neq i} \phi_j^i \hat{Y}_{jt}\right)^2$ is continuous with probability one by inspection. Given these properties and **A5**, $\frac{1}{T_0} \sum\limits_{t=1}^{T_0} \left(Y_{it} - \sum\limits_{j \neq i} \phi_j^i \hat{Y}_{jt}\right)^2$ converges uniformly to $E\left[\frac{1}{T_0} \sum\limits_{t=1}^{T_0} \left(Y_{it} - \sum\limits_{j \neq i} \phi_j^i \hat{Y}_{jt}\right)^2\right]$ by Theorem 2.3 in White and Domowitz (1984).

Equation (14) shows that

$$\left\{ \frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{it} - \sum_{j \neq i} \phi_j^i \hat{Y}_{jt}\right)^2 \right\}$$

is minimized at $\boldsymbol{w}_i$, which is unique by **A6**. Under these conditions, the synthetic control weights are consistent (see Theorem 2.1 of Newey and McFadden (1994)) such that $\widehat{\boldsymbol{w}}_i \xrightarrow{p} \boldsymbol{w}_i$.

Consider unit 1 for $t > T_0$ assuming that **A2**(a) holds. By consistency of $\widehat{\boldsymbol{w}}_1$,

$$\hat{\alpha}_{1t} \xrightarrow{p} \alpha_{1t} + \left(\boldsymbol{\lambda}_t \boldsymbol{\mu}_1 - \boldsymbol{\lambda}_t \sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j\right) + \left(\epsilon_{1t} - \sum_{j=2}^{N} w_j^1 \epsilon_{jt}\right).$$

We know that

$$E\left[\boldsymbol{\lambda}_t \boldsymbol{\mu}_1 - \boldsymbol{\lambda}_t \sum_{j=2}^{N} w_j^1 \boldsymbol{\mu}_j\right] = 0,$$

$$\text{and } E\left[\epsilon_{1t} - \sum_{j=2}^{N} w_j^1 \epsilon_{jt}\right] = 0.$$

Thus, $\hat{\alpha}_{1t}$ is asymptotically unbiased.

Alternatively, consider unit i in which **A2**(b) holds.

$$-\hat{\alpha}_{1t}\hat{w}_1^i \xrightarrow{p} -\alpha_{1t}w_1^i + \left(\boldsymbol{\lambda}_t\boldsymbol{\mu}_i - \boldsymbol{\lambda}_t\sum_{j\neq i} w_j^i\boldsymbol{\mu}_j\right) + \left(\epsilon_{it} - \sum_{j\neq i} w_j^i\epsilon_{jt}\right)$$

The proof follows as before.

$\square$

Additional Figures

Figure A.1: Health Insurance Event Study Estimates

Notes: Using the weights generated from ISCM (all states), corresponding to the last column of Table 5 Panel A, I present the difference in insured rates over time using rates generated from the Current Population Study. Unlike the graphs in Figure 2, "year" in this graphs refer to the calendar year since the CPS Annual Social and Economic Supplement is an annual survey (hence, we should not expect to observe the health insurance coverage effect to occur until 2007).

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program," *Journal of the American Statistical Association*, 2010, *105* (490).

_ , _ , and _ , "Comparative politics and the synthetic control method," *American Journal of Political Science*, 2015, *59* (2), 495–510.

_ and Javier Gardeazabal, "The economic costs of conflict: A case study of the Basque Country," *American Economic Review*, 2003, pp. 113–132.

Acemoglu, Daron, Simon Johnson, Amir Kermani, James Kwak, and Todd Mitton, "The value of connections in turbulent times: Evidence from the United States," *Journal of Financial Economics*, 2016, *121* (2), 368–391.

Adhikari, Bibek and James Alm, "Evaluating the Economic Effects of Flat Tax Reforms Using Synthetic Control Methods," *Southern Economic Journal*, 2016.

_ , Romain A Duval, Bingjie Hu, Prakash Loungani et al., "Can Reform Waves Turn the Tide? Some Case Studies Using the Synthetic Control Method," Technical Report, International Monetary Fund 2016.

Akcigit, Ufuk, Salome Baslandze, and Stefanie Stantcheva, "Taxation and the international mobility of inventors," *The American Economic Review*, 2016, *106* (10), 2930–2981.

Ando, Michihito, "Dreams of urbanization: Quantitative case studies on the local impacts of nuclear power facilities using the synthetic control method," *Journal of Urban Economics*, 2015, *85*, 68–85.

_ and Fredrik Sävje, "Hypothesis Testing with the Synthetic Control Method," 2013.

Athey, Susan and Guido W Imbens, "The state of applied econometrics: Causality and policy evaluation," *Journal of Economic Perspectives*, 2017, *31* (2), 3–32.

_ , Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi, "Matrix completion methods for causal panel data models," *arXiv preprint arXiv:1710.10251*, 2017.

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

Bakirov, Nail Kutluzhanovich and Gábor J Székely, "Student's t-test for Gaussian scale mixtures," *Journal of Mathematical Sciences*, 2006, *139* (3), 6497–6505.

Bauhoff, Sebastian, "The effect of school district nutrition policies on dietary intake and overweight: A synthetic control approach," *Economics & Human Biology*, 2014, *12*, 45–55.

Billmeier, Andreas and Tommaso Nannicini, "Assessing economic liberalization episodes: A synthetic control approach," *Review of Economics and Statistics*, 2013, *95* (3), 983–1001.

Bohn, Sarah, Magnus Lofstrom, and Steven Raphael, "Did the 2007 Legal Arizona Workers Act reduce the state's unauthorized immigrant population?," *Review of Economics and Statistics*, 2014, *96* (2), 258–269.

Canay, Ivan A, Joseph P Romano, and Azeem M Shaikh, "Randomization tests under an approximate symmetry assumption," *Econometrica*, 2017, *85* (3), 1013–1030.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano, "Catastrophic natural disasters and economic growth," *Review of Economics and Statistics*, 2013, *95* (5), 1549–1561.

Chernozhukov, Victor, Kaspar Wuthrich, and Yinchu Zhu, "An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls," *arXiv preprint arXiv:1712.09089*, 2018.

Conley, Timothy G and Christopher R Taber, "Inference with difference in differences with a small number of policy changes," *The Review of Economics and Statistics*, 2011, *93* (1), 113–125.

Courtemanche, Charles J and Daniela Zapata, "Does universal coverage improve health? The Massachusetts experience," *Journal of Policy Analysis and Management*, 2014, *33* (1), 36–69.

Cunningham, Scott and Manisha Shah, "Decriminalizing indoor prostitution: Implications for sexual violence and public health," Technical Report, National Bureau of Economic Research 2014.

Dillender, Marcus O, Carolyn J Heinrich, and Susan N Houseman, "Health insurance reform and part-time work: Evidence from Massachusetts," *Labour Economics*, 2016.

Donohue, John, Abhay Aneja, and Kyle D. Weber, "Do Handguns Make Us Safer? A State-Level Synthetic Controls Analysis of Right-to-Carry Laws," 2015.

Doudchenko, Nikolay and Guido W Imbens, "Balancing, regression, difference-in-differences and synthetic control methods: A synthesis," Technical Report, National Bureau of Economic Research 2016.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler, "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 2017, *132* (1), 435–483.

Eren, Ozkan and Serkan Ozbeklik, "What Do Right-to-Work Laws Do? Evidence from a Synthetic Control Method Analysis," *Journal of Policy Analysis and Management*, 2016, *35* (1), 173–194.

Ferman, Bruno and Cristine Pinto, "Inference in differences-in-differences with few treated groups and heteroskedasticity," 2015.

_ **and** _ , "Revisiting the synthetic control estimator," 2016.

Firpo, Sergio and Vítor Possebom, "Synthetic Control Estimator: A Generalized Inference Procedure and Confidence Sets," *Unpublished manuscript*, 2016.

Fletcher, Jason M, David E Frisvold, and Nathan Tefft, "Non-linear effects of soda taxes on consumption and weight outcomes," *Health Economics*, 2014.

Gobillon, Laurent and Thierry Magnac, "Regional policy evaluation: Interactive fixed effects and synthetic controls," *Review of Economics and Statistics*, 2016, *98* (3), 535–551.

Grier, Kevin and Norman Maynard, "The economic consequences of Hugo Chavez: A synthetic control analysis," *Journal of Economic Behavior & Organization*, 2016, *125*, 1–21.

Hackmann, Martin B, Jonathan T Kolstad, and Amanda E Kowalski, "Adverse selection and an individual mandate: When theory meets practice," *The American Economic Review*, 2015, *105* (3), 1030–1066.

Hahn, Jinyong and Ruoyao Shi, "Synthetic control and inference," *Econometrics*, 2017, *5* (4), 52.

Harrell, Frank E, *Regression modeling strategies, with applications to linear models, survival analysis and logistic regression* 2001.

Hinrichs, Peter, "The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities," *Review of Economics and Statistics*, 2012, *94* (3), 712–722.

Ibragimov, Rustam and Ulrich K Müller, "t-Statistic based correlation and heterogeneity robust inference," *Journal of Business & Economic Statistics*, 2010, *28* (4), 453–468.

Kaul, Ashok, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler, "Synthetic Control Methods: Never Use All Pre-Intervention Outcomes as Economic Predictors," 2015.

Kleven, Henrik Jacobsen, Camille Landais, and Emmanuel Saez, "Taxation and international migration of superstars: Evidence from the European football market," *The American Economic Review*, 2013, *103* (5), 1892–1924.

Kolstad, Jonathan T and Amanda E Kowalski, "The impact of health care reform on hospital and preventive care: evidence from Massachusetts," *Journal of Public Economics*, 2012, *96* (11), 909–929.

Kreif, Noémi, Richard Grieve, Dominik Hangartner, Alex James Turner, Silviya Nikolova, and Matt Sutton, "Examination of the synthetic control method for evaluating health policies with multiple treated units," *Health Economics*, 2015.

Long, Sharon K, "On the road to universal coverage: impacts of reform in Massachusetts at one year," *Health Affairs*, 2008, *27* (4), w270–w284.

_ , Karen Stockley, and Heather Dahlen, "Massachusetts health reforms: uninsurance remains low, self-reported health status improves as state prepares to tackle costs," *Health Affairs*, 2012, pp. 10–1377.

_ , _ , and Kate Willrich Nordahl, "Coverage, access, and affordability under health reform: learning from the Massachusetts model," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 2012, *49* (4), 303–316.

Mideksa, Torben K, "The economic impact of natural resources," *Journal of Environmental Economics and Management*, 2013, *65* (2), 277–289.

Miller, Sarah, "The effect of the Massachusetts reform on health care utilization," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 2012, *49* (4), 317–326.

Montalvo, Jose G, "Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections," *Review of Economics and Statistics*, 2011, *93* (4), 1146–1154.

Munasib, Abdul and Dan S Rickman, "Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis," *Regional Science and Urban Economics*, 2015, *50*, 1–17.

_ and Mouhcine Guettabi, "Florida Stand Your Ground Law and Crime: Did It Make Floridians More Trigger Happy?," *Available at SSRN 2315295*, 2013.

Nannicini, Tommaso and Andreas Billmeier, "Economies in transition: How important is trade openness for growth?," *Oxford Bulletin of Economics and Statistics*, 2011, *73* (3), 287–314.

Newey, Whitney K. and Daniel McFadden, "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 1994, *4*, 2111–2245.

Nonnemaker, James, Mark Engelen, and Daniel Shive, "Are methamphetamine precursor control laws effective tools to fight the methamphetamine epidemic?," *Health Economics*, 2011, *20* (5), 519–531.

Pande, Aakanksha H, Dennis Ross-Degnan, Alan M Zaslavsky, and Joshua A Salomon, "Effects of healthcare reforms on coverage, access, and disparities: quasi-experimental analysis of evidence from Massachusetts," *American Journal of Preventive Medicine*, 2011, *41* (1), 1–8.

Pinotti, Paolo, "The economic costs of organised crime: Evidence from Southern Italy," *The Economic Journal*, 2015, *125* (586), F203–F232.

Powell, David, "Synthetic Control Estimation Beyond Case Studies: Does the Minimum Wage Reduce Employment?," 2016.

Quast, Troy and Fidel Gonzalez, "Sex Work Regulation and Sexually Transmitted Infections in Tijuana, Mexico," *Health Economics*, 2016.

Saunders, Jessica, Russell Lundberg, Anthony A Braga, Greg Ridgeway, and Jeremy Miles, "A Synthetic Control Approach to Evaluating Place-Based Crime Interventions," *Journal of Quantitative Criminology*, 2014, pp. 1–22.

Sommers, Benjamin D, Sharon K Long, and Katherine Baicker, "Changes in Mortality After Massachusetts Health Care Reform: A Quasi-experimental Study," *Annals of Internal Medicine*, 2014, *160* (9), 585–593.

Torsvik, Gaute and Kjell Vaage, "Gatekeeping versus monitoring: Evidence from a case with extended self-reporting of sickness absence," 2014.

Webb, Matthew D, "Reworking wild bootstrap based inference for clustered errors," Technical Report, Queen's Economics Department Working Paper 2013.

White, Halbert and Ian Domowitz, "Nonlinear regression with dependent observations," *Econometrica: Journal of the Econometric Society*, 1984, pp. 143–161.

Zhang, Jun, Chunping Zhong, and Ming Yi, "Did Olympic Games improve air quality in Beijing? Based on the synthetic control method," *Environmental Economics and Policy Studies*, 2016, *18* (1), 21–39.