

Article

Synthetic Control and Inference

Jinyong Hahn ^{1,†} and Ruoyao Shi ^{2,*,†} 

¹ Department of Economics, University of California Los Angeles, 8283 Bunche Hall, Los Angeles, CA 90095, USA; hahn@econ.ucla.edu

² Department of Economics, University of California Riverside, 3136 Sproul Hall, Riverside, CA 92521, USA

* Correspondence: ruoyao.shi@ucr.edu; Tel.: +1-951-827-1494

† These authors contributed equally to this work.

Academic Editors: In Choi and Ryo Okui

Received: 27 October 2016; Accepted: 3 November 2017; Published: 28 November 2017

Abstract: We examine properties of permutation tests in the context of synthetic control. Permutation tests are frequently used methods of inference for synthetic control when the number of potential control units is small. We analyze the permutation tests from a repeated sampling perspective and show that the size of permutation tests may be distorted. Several alternative methods are discussed.

Keywords: synthetic control; permutation test; symmetry

JEL Classification: C12

1. Introduction

Synthetic control method, proposed and discussed by [Abadie and Gardeazabal \(2003\)](#) and [Abadie et al. \(2010\)](#), is a very useful way of conducting comparative studies when exact matches are unavailable. Estimation of treatment effects usually takes the form of comparing outcomes between the treated unit and the control unit. Common sense suggests that, for the comparison to be meaningful, the control unit needs to be similar to the treated unit in the absence of the treatment in various dimensions. Such a requirement may not be satisfied in many observational studies. In some cases, availability of panel data makes such comparisons reasonable, the difference-in-differences method being a very well-known example. The difference-in-differences method requires a very specific set of assumptions, i.e., the common trend assumption, which may not be plausible for many applications. The synthetic control method offers a sensible generalization of the difference-in-differences. The synthetic control is a linear combination of the potential control outcomes, where the weights are manufactured by analyzing the pre-intervention outcomes.

For the purpose of statistical inference with synthetic control, i.e., confidence interval and hypothesis testing, various versions of placebo tests are often adopted. The idea underlying the placebo tests is the usual permutation tests, where the critical value of a test statistic is computed under all possible permutations of the “treatment” assignments in the control units.

The idea of permutation test is very intuitive and attractive. Applying the synthetic control method to every potential control unit presumably allows researchers to assess the distribution of a test statistic under the null hypothesis of no treatment effects, and the inference is seemingly exact in the sense that the burden of asymptotic approximation can be obviated.

The purpose of this paper is very specific. We ask whether the permutation test is a reasonable idea in the context of the synthetic control method, and argue that the intuitive appeal of the permutation test is misplaced. The validity of permutation tests usually requires certain symmetry assumption, which is often violated in the context of synthetic control studies. Using Monte Carlo simulations,

we document the size distortion of the permutation tests. We also discuss a few alternative methods of inference.

Alberto Abadie kindly pointed out that the placebo test in synthetic control is often based on randomization inference idea, under which the symmetry restriction is built-in, while our analysis is predicated from the usual random sampling perspective, which leads to the violation of symmetry. This perspective is shared with an anonymous referee, who notes that (i) the synthetic control literature uses permutation tests in the context of design-based inference, and, as such, the permutation tests have exact size; (ii) the present article shows that permutation tests may not have correct size under a different mode of inference based on repeated sampling, although interpreting the permutation tests in the previous literature as tests based on repeated sampling would be incorrect; and (iii) the present article also proposes some alternatives that are valid in a repeated sampling setting. It would be useful to understand the exact mechanism through which the difference between the two perspectives manifests itself. The same referee points out that the present paper adopts a setting where $T_0 \rightarrow \infty$, while the original literature assumes fixed T_0 .

2. Placebo Test and Synthetic Control

In this section, we provide a brief discussion of the placebo test in the context of the synthetic control method. We begin with an overview of the synthetic control, borrowing heavily from discussions in [Abadie et al. \(2010\)](#) and [Doudchenko and Imbens \(2016\)](#). We then move on to describe the placebo test, and point out the importance of the symmetry assumption. We argue that the symmetry assumption is violated in general for placebo tests using linear combinations of outcomes, such as synthetic control. We conclude this section that such violation should be expected in general even when a normalized version of the test statistic is adopted.

We start with the overview of the synthetic control method. Consider a panel data with $J + 1$ cross sectional units observed over the time periods $t = 1, \dots, T$. Units $j = 1, \dots, J$ are the control units that receive the treatment in none of the time periods. The unit $j = 0$ receives no treatment in periods $1, \dots, T_0$, and receives active treatment in time periods $t = T_0 + 1, \dots, T$. For simplicity, we will often assume that $T = T_0 + 1$. The outcome variable $Y_{j,t}$ is such that $Y_{j,t} = Y_{j,t}(1)$ if the j th unit receives treatment in time t , and $Y_{j,t} = Y_{j,t}(0)$ otherwise. Obviously,

$$\begin{aligned} Y_{j,t} &= Y_{j,t}(0), & j = 1, \dots, J; t = 1, \dots, T, \\ Y_{0,t} &= Y_{0,t}(0), & t = 1, \dots, T_0, \\ Y_{0,t} &= Y_{0,t}(1), & t = T_0 + 1, \dots, T. \end{aligned}$$

The idea underlying the synthetic control is that if there were some weights¹ $\hat{\omega}_1, \dots, \hat{\omega}_J$ such that

$$Y_{0,t} \approx \sum_{j=1}^J \hat{\omega}_j Y_{j,t} \tag{1}$$

during the pre-intervention periods ($t = 1, \dots, T_0$). Then, $\sum_{j=1}^J \hat{\omega}_j Y_{j,t}$ can be used as a (synthetic) control for $Y_{0,t}$ during the post-intervention periods ($t = T_0 + 1, \dots, T$). [Abadie et al. \(2010\)](#) and [Doudchenko and Imbens \(2016\)](#) discuss various methods of finding the $\hat{\omega}$'s so that the requirement in Equation (1) is satisfied. We analyze the weights and the nature of approximation from the asymptotic

¹ [Doudchenko and Imbens \(2016\)](#) also consider a slightly more general requirement $Y_{0,t} \approx \alpha + \sum_{j=1}^J w_j Y_{j,t}$. This is a sensible way to enhance accuracy of synthetic control viewed as a point estimator. It also provides a link to the difference-in-differences estimator. Because our focus is on inferential aspects of the problem, we simplify notation and analysis by abstracting away from the intercept term.

perspective where $T_0 \rightarrow \infty$. Note that a special case of the estimator discussed by [Abadie et al. \(2010, p. 496\)](#) solves

$$\min_{\omega_1, \dots, \omega_J} \left(\bar{Y}_0 - \sum_{j=1}^J \omega_j \bar{Y}_j \right)^2. \tag{2}$$

Under our interpretation, $\bar{Y}_j = T_0^{-1} \sum_{t=1}^{T_0} Y_{j,t}$ above is an estimator of $E[Y_{j,t}]$. Consider the linear factor structure² as in [Abadie et al. \(2010\)](#):

$$Y_{j,t} = Y_{j,t}(0) = \alpha_j + \theta_t + \gamma_j' \delta_t + \epsilon_{j,t}, \quad t = 1, \dots, T_0. \tag{3}$$

Suppose that θ_t , δ_t , and $\epsilon_{j,t}$ satisfy strict stationarity. Without loss of generality, we also assume that $E[\delta_t] = 0$ and $E[\epsilon_{j,t}] = 0$. We would have $\bar{Y}_j \rightarrow \alpha_j + E[\theta_t]$ in probability as $T_0 \rightarrow \infty$. Assuming that ω_{*j} satisfies

$$\alpha_0 + E[\theta_t] = \sum_{j=1}^J \omega_{*j} (\alpha_j + E[\theta_t]), \tag{4}$$

we can understand that the population version of the synthetic control $\sum_{j=1}^J \omega_{*j} Y_{j,t}$ is such that the difference $Y_{0,T_0+1}(0) - \sum_{j=1}^J \omega_{*j} Y_{j,T_0+1}(0)$ is designed to have a mean zero.

Our $T_0 \rightarrow \infty$ asymptotic interpretation is not the only possible one. [Doudchenko and Imbens \(2016\)](#) provide an in-depth analysis of many possible methods. Our interpretation, however, is helpful for two reasons. First, it makes a concrete interpretation of $\hat{\omega}$ s as estimates of some pseudo-parameter, say ω_{*} 's, along with analytic expressions of the ω_{*} 's, which makes it easy to understand the potential pitfalls of permutation methods afterwards. Second, it helps us to motivate alternative methods of inference exploiting time series variation.

We now discuss how placebo tests can be used in the context of synthetic control. For this purpose, we first present a summary of the placebo tests/permutation tests. The tests are motivated to deal with the case where the number of the treated is small and the number of controls is relatively large. In order to focus on the salient feature of the tests, we will consider an extreme case and assume that there is only one treated unit.

The basic intuition underlying the general placebo test can be gleaned by examining a standard textbook case of randomized treatments. Suppose that there is cross sectional data with $J + 1$ units, where the units $j = 1, \dots, J$ are the control units and the unit $j = 0$ receives the active treatment. A reasonable estimator of the treatment effect is the difference $Y_0 - \bar{Y}$, where Y_0 is the outcome of the unit $j = 0$, and $\bar{Y} = J^{-1} \sum_{j=1}^J Y_j$ denotes the average of the outcomes of the controls. Suppose that we are interested in testing whether the treatment had impact. Given that there is only one treated unit, the standard t -test comparing the difference of the mean outcomes is not applicable. On the other hand, common sense suggests that we may implement such a test by “assigning” each control unit to fictitious treatment. More precisely, one can estimate the empirical distribution of $Y_k - (J - 1)^{-1} \sum_{j \neq k} Y_j$ for $k = 1, \dots, J$, and use it as if it were the distribution of the treatment effect under the null hypothesis.³

Implementation of the placebo test with synthetic control requires a bit more notation. First let $\hat{\omega} = (\hat{\omega}_1, \dots, \hat{\omega}_J)'$ denote the estimator of $\omega_{*} = (\omega_{*1}, \dots, \omega_{*J})'$. Although we will use the method of exact balancing later in our Monte Carlo simulations, we do not need to restrict ourselves to this particular estimator. For now, we can view $\hat{\omega}$ as an output from a blackbox and let ω_{*} denote its

² Using the notation consistent with this paper, Equation (1) in [Abadie et al. \(2010\)](#) takes the form $Y_{j,t}(0) = \lambda_t + (\theta_t - \lambda_t)' Z_j + \gamma_j' \delta_t + \epsilon_{j,t}$, so the factor structure in Equation (3) of this paper is a special case of Equation (1) in [Abadie et al. \(2010\)](#), where $Z_j = 1$, $\delta_{1t} = 1$ and $\gamma_{1j} = \alpha_j$, i.e., it is a special case where the Z_j does not exist and the first element of δ_t is time invariant.

³ [Conley and Taber \(2011\)](#), who proposed a similar test, cite [Bertrand et al. \(2004\)](#) when they discuss placebo tests. [Abadie et al. \(2010\)](#) reference many other papers that precede [Bertrand et al. \(2004\)](#).

probability limit as $T_0 \rightarrow \infty$. Second, let $\hat{\omega}^{(-k)}$ denote the outcome of the same blackbox except that we use the k th unit as the outcome of the treated unit, and $Y_{j,t}$ with $j \neq k$ as our control units. The placebo test then uses the empirical distribution of $Y_{k,T_0+1} - \sum_{j \neq k} \hat{\omega}_j^{(-k)} Y_{j,T_0+1}$ for $k = 1, \dots, J$ as if it were the distribution of the treatment effect under the null hypothesis of no treatment effect. If the estimated effect $Y_{0,T_0+1} - \sum_{j=1}^J \hat{\omega}_j Y_{j,T_0+1}$ belongs to the extreme tails of the empirical distribution, it is understood to be the evidence that the null hypothesis is incorrect.

In order to understand the size property of the placebo test, it helps to recall that the placebo test is a version of the permutation test, which requires for its validity what may be called the symmetry assumption. For review of this property, we will borrow the short discussion in [Canay et al. \(2017\)](#).⁴ Suppose that a researcher observes a vector of observations X , whose joint distribution is P . The objective is to test whether $P \in \mathbf{P}_0$, where \mathbf{P}_0 is a collection of probability distributions such that the distribution of X is equal to that of gX for every g in \mathbf{G} , where \mathbf{G} is a finite collection of transformations. The permutation test has the exact size if, for the test statistic $T(X)$, the critical value is taken from the distribution of $T(gX)$ for every g in \mathbf{G} . In the context of the placebo test above, one can understand X to be the vector (Y_1, \dots, Y_J) , and gX to be the permutation of the Y s.

We note that the symmetry is not mathematically obvious in the context of synthetic control. In order for the permutation test to be valid, it is necessary for the distribution of $Y_{0,T_0+1} - \sum_{j=1}^J \omega_{*j} Y_{j,T_0+1}$ and those of $Y_{k,T_0+1} - \sum_{j \neq k} \omega_{*j}^{(-k)} Y_{j,T_0+1}$ for $k = 1, \dots, J$ to be identical. Even for the relatively simple model in Equation (3), the nature of the synthetic control is such that the symmetry does not naturally follow. Using the restriction in Equation (4), we may write

$$Y_{0,T} - \sum_{j=1}^J \omega_{*j} Y_{j,T} = (\theta_T - E[\theta_T]) \left(1 - \sum_{j=1}^J \omega_{*j}\right) + \left(\gamma_0 - \sum_{j=1}^J \omega_{*j} \gamma_j\right) \delta_T + \left(\epsilon_{0,T} - \sum_{j=1}^J \omega_{*j} \epsilon_{j,T}\right). \quad (5)$$

Even if the first two terms on the right-hand side of Equation (5) were identically equal to zero over the permutations, we believe that the third term is not likely to satisfy the symmetry property. This is because we believe that under the further restriction that the ϵ 's have a finite variance, the term can be symmetric only when they are normally distributed.

We show that normality is necessary if the distribution of the error term $\epsilon_{0,T} - \omega'_{*} \epsilon_T$ in Equation (5), where $\epsilon_T = (\epsilon_{1,T}, \dots, \epsilon_{J,T})'$, is to be symmetric up to normalization.⁵ Suppose that $\epsilon_{0,T}, \dots, \epsilon_{J,T}$ are i.i.d., and their common distribution is such that the variance is finite and the characteristic function does not disappear. If ω is a nontrivial function of α s and γ s, then symmetry over the permutations requires that the marginal distributions of $\epsilon_{k,T} - \sum_{j \neq k} \omega_j^{(-k)} \epsilon_{j,T}$ for $k = 0, \dots, J$ should remain invariant over all possible $\omega^{(-k)}$ s. Without loss of generality, we can focus on the distribution of $\epsilon_{0,T} - \omega' \epsilon_T$, and conclude that the symmetry requires that there exists a random variable Y such that the distribution of $\epsilon_{0,T} - \omega' \epsilon_T$ is the same as that of cY for some scalar c . Because the standard deviation of $\epsilon_{0,T} - \omega' \epsilon_T$ is proportional to $\sqrt{1 + \omega' \omega}$, we may without loss of generality take $c = \sqrt{1 + \omega' \omega}$. This implies that the distribution of $\omega' \epsilon_T$ only depends on $\omega' \omega$. In other words, for $\omega \neq \tilde{\omega}$ such that $\omega' \omega = \tilde{\omega}' \tilde{\omega}$, the distribution of $\omega' \epsilon_T$ is identical to that of $\tilde{\omega}' \epsilon_T$. In particular, let all components of $\tilde{\omega}$ be zero except for the first one. Then, the distribution of $\omega' \epsilon_T$ is identical to that of $\tilde{\omega}' \epsilon_T = \tilde{\omega}_1 \epsilon_{1,T} = \sqrt{\tilde{\omega}' \tilde{\omega}} \epsilon_{1,T}$. This implies that $\epsilon_{j,T}$ should have a stable distribution.⁶ Because the only stable distribution with a finite variance is the normal distribution, we should conclude that normality is a necessary condition of the symmetry (up to normalization). Note that the third term in Equation (5) arises in an ideal

⁴ The same test was first discussed by [Hoeffding \(1952\)](#), which is a generalization of the randomization test proposed by [Fisher \(1949\)](#).

⁵ We are using the fact that the symmetry implies the equality of marginal distributions, and therefore, the lack of equality of marginal distributions is a sufficient condition for violation of symmetry.

⁶ See [Nolan \(2015\)](#), or [Wikipedia Contributors \(2017\)](#).

situation where the weights ω do not need to be estimated and the first two terms completely disappear. Our analysis suggests that even if we normalize the third term by its standard deviation, the symmetry requires normal distribution. The necessity of normality assumption is about *any* linear combination so it applies a fortiori to synthetic control.

3. Monte Carlo

The discussion at the end of the previous section casts doubt on the placebo test, even for the simple case where the first two terms in Equation (5) can be ignored. In order to understand the roles that the first two terms may play, we adopt Monte Carlo simulations. We try to find data generating processes (DGPs hereafter) that generate a large amount of size distortions. This is helpful in understanding the potential problem of the placebo test from the uniformity perspective; after all, the mathematical definition of the “size” of a test is the maximum probability of rejection under the null, and here the null hypothesis is a composite hypothesis where the only requirement on the DGP is that the treatment effect is zero, which allows many possibilities on the terms in Equation (5). For this purpose, we found it most convenient to work with the first two terms in Equation (5), although we acknowledge that there may be other important sources of size distortion that we have not explored. Since the last paragraph of Section 2 showed that normalization does not abate the symmetry requirement, we examine the importance of the first two terms in Equation (5) using a more natural statistic. The version of the synthetic control that we use in the Monte Carlo is the method of exact balancing, the population version of which minimizes $\sum_{j=1}^J \omega_j^2$ subject to $E[Y_{0,t}(0)] = \sum_{j=1}^J \omega_j E[Y_{j,t}(0)]$ and $1 = \sum_{j=1}^J \omega_j$.⁷

The method of exact balancing may not be an ideal version of the synthetic control, but it reflects a certain ambiguity in the method of synthetic control. In the factor model in Equation (3), it is impossible to find weights ω such that $Y_{0,t} = \sum_{j=1}^J \omega_j Y_{j,t}$ for every $t = 1, \dots, T_0$, if T_0 is large enough, as long as $\epsilon_{j,t}$ is continuously distributed. In other words, the condition (2) in Abadie et al. (2010) is incompatible with the factor model unless $\text{Var}(\epsilon_{j,t}) = 0$. The assumption $\text{Var}(\epsilon_{j,t}) = 0$ has at least two implications.⁸ First, the weights ω_j can be estimated without error with sufficiently large T_0 . Second, the distribution of the permutation test would have the point mass at zero, and as such, there is no reason to conduct any test. Both implications are questionable. In any case, under the assumption $\text{Var}(\epsilon_{j,t}) = 0$, the weights can be estimated (without error) by the method of least squares that minimizes $\sum_{t=1}^{T_0} \left(Y_{0,t} - \sum_{j=1}^J \omega_j Y_{j,t} \right)^2$. If the assumption $\text{Var}(\epsilon_{j,t}) = 0$ is violated, the method of least squares would be subject to a version of measurement error problem; the true regressor there is $\alpha_j + \theta_t + \gamma_j' \delta_t$ in Equation (3), and the $Y_{j,t}$ plays the role of a regressor with measurement error $\epsilon_{j,t}$.⁹ Note that such a problem is avoided by the method of exact balancing.

We consider the method of exact balancing in this section not because it is necessarily an ideal version of the synthetic control, but because it is a convenient way of examining the impact of the first two terms in Equation (5). As mentioned at the beginning of this section, our analysis at the end of the previous section suggests that the placebo test may have a problem even when these two terms are dismissed, and the purpose of our Monte Carlo exercise is to focus on the potential impact of these two terms.

⁷ Abadie et al. (2010) also impose the positivity restriction, i.e., $\omega_j \geq 0$ for all j .

⁸ It is straightforward to prove that under stationarity assumption, the only model that allows the synthetic controls to trace the trajectory of the outcome for the treated (i.e., $Y_{0,t} = \sum_{j=1}^J \omega_j Y_{j,t}$ for some ω) is a linear factor model with $\text{Var}(\epsilon_{j,t}) = 0$.

⁹ See Ferman and Pinto (2017) for related discussion on the bias of the synthetic control estimator.

For our Monte Carlo analysis, we adopted a simplified version of the factor model in Equation (3) such that (i) $\theta_t \sim \mathcal{N}(0, 1)$; (ii) δ_t is a scalar; (iii) $\delta_t \sim \mathcal{N}(0, \sigma_\delta^2)$; (iv) $\epsilon_{j,t} \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is i.i.d. over j and t . In matrix notation, our estimator $\hat{\omega}$ solves

$$\min \omega' \omega \quad \text{s.t.} \quad \text{(i) } \bar{Y}' \omega = \bar{Y}_0, \text{ (ii) } \ell' \omega = 1, \tag{6}$$

where ℓ is a vector of ones. Because $E[Y_{j,t}] = \alpha_j$, we can see that the population counterpart ω_* solves

$$\min \omega' \omega \quad \text{s.t.} \quad \text{(i) } \alpha' \omega = \alpha_0, \text{ (ii) } \ell' \omega = 1,$$

where $\alpha = (\alpha_1, \dots, \alpha_J)'$.

We now write

$$Y_{0,T} - \sum_{j=1}^J \hat{\omega}_j Y_{j,T} = A + B + C + D,$$

where

$$\begin{aligned} A &= \sum_{j=1}^J (\omega_{*j} - \hat{\omega}_j) \alpha_j \\ B &= \underbrace{\theta_T \left(1 - \sum_{j=1}^J \omega_{*j} \right)}_{B(i)} + \underbrace{\theta_T \sum_{j=1}^J (\omega_{*j} - \hat{\omega}_j)}_{B(ii)} \\ C &= \underbrace{\left(\gamma_0 - \sum_{j=1}^J \omega_{*j} \gamma_j \right)'}_{C(i)} \delta_T + \underbrace{\left(\sum_{j=1}^J (\omega_{*j} - \hat{\omega}_j) \gamma_j \right)'}_{C(ii)} \delta_T \\ D &= \underbrace{\left(\epsilon_{0,T} - \sum_{j=1}^J \omega_{*j} \epsilon_{j,T} \right)}_{D(i)} + \underbrace{\left(\sum_{j=1}^J (\omega_{*j} - \hat{\omega}_j) \epsilon_{j,T} \right)}_{D(ii)}. \end{aligned}$$

Note that the term B(i) is equal to 0 by design here, although it can be in principle different from 0 depending on the DGP and the estimator chosen. We speculate that the placebo test is used in the hope that (a) $Y_{0,T} - \sum_{j=1}^J \hat{\omega}_j Y_{j,T}$ is dominated by the term D(i) above; (b) the four terms A, B(ii), C(ii) and D(ii) above, which reflect the noise of estimating ω_* by $\hat{\omega}$, are ignorable; and (c) the two terms C(i) and D(i) more or less satisfy the symmetry property.

We argued in the previous section that the term D(i) is likely to violate the symmetry property. In order to assess the impacts of other terms, we consider the following variations in DGPs:

1. Vary the values of α 's such that (a) none of the components of ω_* dominates; (b) only two of the elements are non-zero.
2. Vary the values of γ 's such that the unbalanced unobservable factors C(i) (a) disappear; and (b) are present.
3. Vary T_0 such that the estimation errors in the weights are (a) prominent; and (b) negligible.

Combinations of the first two variations give us four different DGPs, shown as DGP No. 1 to No. 4 in Table 1.

Table 1. Data Generating Processes (DGPs) that generate size distortion.

DGP No.	α 's	γ 's	Variations
1	$\alpha_0 = (2J + 1)/3, \alpha_1 = 1, \dots, \alpha_J = J$	$\gamma_0 = \gamma_1 = \dots = \gamma_J = 0$	1(a), 2(a)
2	$\alpha_0 = (2J + 1)/3, \alpha_1 = 1, \dots, \alpha_J = J$	$\gamma_0 = 2, \gamma_1 = \dots = \gamma_J = 1$	1(a), 2(b)
3	$\alpha_0 = 5/3, \alpha_1 = 1, \alpha_2 = 2, \alpha_3 = \dots = \alpha_J = 0$	$\gamma_0 = \gamma_1 = \dots = \gamma_J = 0$	1(b), 2(a)
4	$\alpha_0 = 5/3, \alpha_1 = 1, \alpha_2 = 2, \alpha_3 = \dots = \alpha_J = 0$	$\gamma_0 = 2, \gamma_1 = \dots = \gamma_J = 1$	1(b), 2(b)

We considered two versions of the placebo tests: the first one is what might be called a feasible version of the test. Formally, for $j = 0, 1, \dots, J$, let Y_j be a $T \times 1$ vector of outcomes for the j th control unit, let $Y = (Y_1, \dots, Y_J)$, and let Y_{-j} be a $T \times (J - 1)$ matrix that deletes the j th column from Y . Then,

$$S(Y_0, Y) \equiv Y_{0,T} - \sum_{j=1}^J \hat{\omega}_{*j} Y_{j,T}. \tag{7}$$

Similar to Equation (6), define the leave-one-out synthetic control weights $\hat{\omega}_{-j}$ for the j th control unit as a solution to

$$\min \omega' \omega \quad \text{s.t.} \quad \text{(i) } \bar{Y}'_{-j} \omega = \bar{Y}_j, \text{ (ii) } \ell' \omega = 1, \tag{8}$$

where \bar{Y}_{-j} is to delete the j th element from \bar{Y} . We likewise define the population counterpart ω_{*-j} as a solution to

$$\min \omega' \omega \quad \text{s.t.} \quad \text{(i) } \alpha'_{-j} \omega = \alpha_j, \text{ (ii) } \ell' \omega = 1.$$

For $j = 1, \dots, J$ and $k \neq j$, let $\hat{\omega}_{-j,k}$ be the element in $\hat{\omega}_{-j}$ that corresponds to the k th control unit. In addition, define $\hat{\omega}_{-j,j} \equiv 0$ for $j = 1, \dots, J$. Then, for $j = 1, \dots, J$, we can compute

$$S(Y_j, Y_{-j}) \equiv Y_{j,T} - \sum_{k=1}^J \hat{\omega}_{-j,k} Y_{k,T}. \tag{9}$$

Let $S^{(1)}, \dots, S^{(J)}$ be the order statistics of $S(Y_j, Y_{-j})$'s. We reject H_0 if $S(Y_0, Y) > S^{(J(1-\frac{\alpha}{2}))}$ or $S(Y_0, Y) < S^{(\frac{J\alpha}{2})}$.

The second test is an infeasible version of the test, which is identical to the first test, except that we use the true value of ω_* , i.e.,

$$S^{true}(Y_0, Y) \equiv Y_{0,T} - \sum_{j=1}^J \omega_{*j} Y_{j,T},$$

$$S^{true}(Y_j, Y_{-j}) \equiv Y_{j,T} - \sum_{k \neq j}^J \omega_{*-j,k} Y_{k,T},$$

and we reject H_0 if $S^{true}(Y_0, Y) > S^{true, (J(1-\frac{\alpha}{2}))}$ or $S^{true}(Y_0, Y) < S^{true, (\frac{J\alpha}{2})}$.

For each DGP, we try $T_0 \in \{40, 80, 400, 800\}$, $J \in \{20, 40, 80\}$ and $\sigma_c^2 = 0.1$. For all designs, we set the level of the tests to be $\alpha = 10\%$, and the number of Monte Carlo runs to be 1000.

The results are summarized in Table 2.¹⁰ We see size distortions in Table 2, especially DGP No. 2 and No. 4. The size distortion there cannot be attributed to the noise of estimating ω . First, the problem persists even as T_0 approaches unrealistically large values. Second, the size distortion is similar over the feasible and infeasible versions of the test. We suspect that the problem is a fundamental problem

¹⁰ We set $\theta_t \sim N(0, 1)$ in Table 2. We also considered the case where $\theta_t = 0$. Although the results for this case are not reported here in the paper, they were qualitatively similar to the $\theta_t \sim N(0, 1)$ case. They are available upon request. (When the adding-up constraint was imposed, the two cases gave the same results. Without the adding-up constraint, these two specifications give slightly different results.)

that may have something to do with the violation of symmetry. (An anonymous referee pointed out that DGPs No. 2 and No. 4 cannot produce synthetic controls that approximate the trajectory of the outcome for the treated, and that synthetic controls should not be applied in those settings.)

Table 2. Null rejection rates of permutation tests.

DGP	J	σ_δ^2	Estimated Weights				True Weights				
			$T_0 = 40$	$T_0 = 80$	$T_0 = 400$	$T_0 = 800$	$T_0 = 40$	$T_0 = 80$	$T_0 = 400$	$T_0 = 800$	
1	J = 20	0.1	0.076	0.067	0.082	0.090	0.074	0.070	0.079	0.091	
		1	0.076	0.067	0.082	0.090	0.074	0.070	0.079	0.091	
		10	0.076	0.067	0.082	0.090	0.074	0.070	0.079	0.091	
	J = 40	0.1	0.079	0.080	0.080	0.093	0.079	0.090	0.084	0.096	
		1	0.079	0.080	0.080	0.093	0.079	0.090	0.084	0.096	
		10	0.079	0.080	0.080	0.093	0.079	0.090	0.084	0.096	
	J = 80	0.1	0.100	0.087	0.113	0.083	0.094	0.087	0.114	0.085	
		1	0.100	0.087	0.113	0.083	0.094	0.087	0.114	0.085	
		10	0.100	0.087	0.113	0.083	0.094	0.087	0.114	0.085	
	2	J = 20	0.1	0.193	0.178	0.175	0.190	0.174	0.176	0.174	0.190
			1	0.568	0.530	0.559	0.529	0.557	0.518	0.561	0.531
			10	0.838	0.843	0.837	0.837	0.839	0.828	0.838	0.838
J = 40		0.1	0.193	0.219	0.198	0.208	0.200	0.213	0.193	0.207	
		1	0.561	0.623	0.586	0.589	0.556	0.593	0.588	0.583	
		10	0.809	0.853	0.863	0.844	0.832	0.858	0.854	0.846	
J = 80		0.1	0.238	0.253	0.245	0.230	0.239	0.245	0.241	0.233	
		1	0.617	0.623	0.631	0.619	0.610	0.629	0.624	0.622	
		10	0.865	0.870	0.870	0.866	0.861	0.894	0.867	0.868	
3		J = 20	0.1	0.115	0.123	0.110	0.125	0.114	0.113	0.109	0.123
			1	0.115	0.123	0.110	0.125	0.114	0.113	0.109	0.123
			10	0.115	0.123	0.110	0.125	0.114	0.113	0.109	0.123
	J = 40	0.1	0.152	0.151	0.149	0.153	0.153	0.154	0.146	0.154	
		1	0.152	0.151	0.149	0.153	0.153	0.154	0.146	0.154	
		10	0.152	0.151	0.149	0.153	0.153	0.154	0.146	0.154	
	J = 80	0.1	0.180	0.168	0.173	0.153	0.178	0.164	0.166	0.153	
		1	0.180	0.168	0.173	0.153	0.178	0.164	0.166	0.153	
		10	0.180	0.168	0.173	0.153	0.178	0.164	0.166	0.153	
	4	J = 20	0.1	0.201	0.219	0.193	0.214	0.207	0.213	0.195	0.209
			1	0.536	0.518	0.525	0.522	0.533	0.520	0.525	0.520
			10	0.837	0.827	0.817	0.814	0.841	0.809	0.824	0.812
J = 40		0.1	0.238	0.253	0.249	0.253	0.241	0.268	0.239	0.249	
		1	0.549	0.597	0.583	0.579	0.552	0.576	0.583	0.582	
		10	0.809	0.842	0.856	0.852	0.828	0.843	0.859	0.848	
J = 80		0.1	0.287	0.293	0.300	0.275	0.279	0.301	0.297	0.278	
		1	0.610	0.623	0.655	0.629	0.608	0.625	0.653	0.634	
		10	0.861	0.871	0.874	0.860	0.866	0.880	0.865	0.866	

Our Monte Carlo analysis indicates that the placebo test does have the size distortion problem. The results in Table 2 suggest that the size problem is potentially bigger in DGPs No. 2 and No. 4. DGPs No. 2 and No. 4 differ from No. 1 and No. 3 in that the γ 's are nonzero and the aggregate shock δ_t plays a role as a consequence. Therefore, it is of interest to investigate further sources of asymmetry. For this purpose, we revisit the decomposition in Equation (5) of $Y_{0,T_0+1}(0) - \sum_{j=1}^J \omega_{*j} Y_{j,T_0+1}(0)$, assuming that the first and second terms in the factor model in Equation (3) are not present.¹¹

¹¹ This can be done by assuming that $\alpha_0 = \sum_{j=1}^J \omega_j \alpha_j$ and $\theta_t = 0$.

$$Y_{0,T_0+1}(0) - \sum_{j=1}^J \omega_{*j} Y_{j,T_0+1}(0) = \left(\gamma_0 - \sum_{j=1}^J \omega_{*j} \gamma_j \right)' \delta_{T_0+1} + \left(\epsilon_{0,T_0+1} - \sum_{j=1}^J \omega_{*j} \epsilon_{j,T_0+1} \right).$$

This implies that the variance of $Y_{0,T_0+1}(0) - \sum_{j=1}^J \omega_{*j} Y_{j,T_0+1}(0)$ can be written as

$$\left(\gamma_0 - \sum_{j=1}^J \omega_{*j} \gamma_j \right)' \Sigma(\delta_t) \left(\gamma_0 - \sum_{j=1}^J \omega_{*j} \gamma_j \right) + \text{var}(\epsilon_{j,t}) \left(1 + \sum_{j=1}^J \omega_{*j}^2 \right)$$

under the assumptions of the DGPs, where $\Sigma(\delta_t)$ is the covariance matrix of the vector δ_t . Likewise, the variances of the permutation statistics are

$$\left(\gamma_j - \sum_{k \neq j} \omega_{*-j,k} \gamma_k \right)' \Sigma(\delta_t) \left(\gamma_j - \sum_{k \neq j} \omega_{*-j,k} \gamma_k \right) + \text{var}(\epsilon_{j,t}) \left(1 + \sum_{k \neq j} \omega_{*-j,k}^2 \right).$$

Depending on the relative magnitudes of γ 's, we can easily construct examples that violate the symmetry, such as DGPs No. 2 and No. 4. As of now, it is not clear to us whether there is another venue (other than the variation in the size of γ), which leads to a violation of the symmetry.

4. Possible Alternatives to Placebo Tests

If we take the time series asymptotics ($T_0 \rightarrow \infty$) seriously, the problem can be avoided by using the same idea as in Andrews (2003). The hypothesis of no treatment effects can be understood to be a hypothesis of stationarity of the time series $W_t \equiv Y_{0,t} - \sum_{j=1}^J \omega_{*j} Y_{j,t}$. In particular, the researcher is interested in whether the distribution of W_{T_0+1}, \dots, W_T is the same as that of W_1, \dots, W_{T_0} , for which Andrews (2003)'s test is well-suited. In the simple case that we consider where $T = T_0 + 1$, one rejects the null if W_{T_0+1} belongs to the extreme tails of the empirical distribution of W_1, \dots, W_{T_0} . We conducted Monte Carlo simulations for all the DGPs considered in the previous section, and verified that Andrews (2003)'s test suffered no size distortion.¹² Andrews (2003)'s test is geared for application in time series, and as such, robust to certain heteroscedasticity. If the variances of $\epsilon_{j,t}$ in Equation (3) were different across j s, most of the available methods exploiting cross sectional variation may need to be used with caution, as noted by Ferman and Pinto (2017). Andrews (2003)'s end-of-sample instability test being a test of stationarity of $Y_{0,t} - \sum_{j=1}^J \omega_{*j} Y_{j,t}$, its validity does not depend on whether the $\epsilon_{j,t}$'s have identical variances or not. The usefulness of Andrews (2003)'s test in this context was recognized earlier by Ferman and Pinto (2017).

Andrews (2003)'s test utilizes time series variation seriously. When T_0 is relatively small, perhaps the researcher would like to have a procedure that is based on cross sectional variation. If the factor structure is taken seriously and if the number of factors is a priori known, we can produce such a procedure by combining the ideas in Conley and Taber (2011) and Holtz-Eakin et al. (1988). For simplicity, assume that the model is given by

$$Y_{j,t} = Y_{j,t}(0) = \gamma_j \delta_t + x'_{j,t} \beta + \epsilon_{j,t} \quad t = 1, 2; j = 1, \dots, J,$$

$$Y_{0,t}(0) = \gamma_0 \delta_t + x'_{0,t} \beta + \epsilon_{0,t},$$

¹² The results are available upon request.

where we normalize $\delta_1 = 1$. Let $Y_{0,1} = Y_{0,1}(0)$ and $Y_{0,2} = Y_{0,2}(1) = Y_{0,2}(0) + \eta$. This is a case where $T_0 = 1$, and $T = 2$. We then have

$$\frac{Y_{j,2}}{\delta_2} - Y_{j,1} = \left(\frac{x_{j,2}}{\delta_2} - x_{j,1} \right)' \beta + \frac{\epsilon_{j,2}}{\delta_2} - \epsilon_{j,1} \quad t = 1, 2; j = 1, \dots, J.$$

Under strict exogeneity assumption on x 's, we can consistently estimate (β, δ_2) as $J \rightarrow \infty$ by using the control group. Now, assume that $(\epsilon_{j,1}, \epsilon_{j,2})$ $j = 0, 1, 2, \dots$ are i.i.d., which would imply

$$\frac{\epsilon_{j,2}}{\delta_2} - \epsilon_{j,1} = \frac{Y_{j,2} - x'_{j,2}\beta}{\delta_2} - (Y_{j,1} - x'_{j,1}\beta)$$

are i.i.d. A simple modification of [Conley and Taber \(2011\)](#)'s argument establishes that the distribution of $\frac{\epsilon_{j,2}}{\delta_2} - \epsilon_{j,1}$ can be consistently estimated by the empirical distribution of

$$\frac{Y_{j,2} - x'_{j,2}\hat{\beta}}{\hat{\delta}_2} - (Y_{j,1} - x'_{j,1}\hat{\beta}) \quad j = 1, \dots, J,$$

where $(\hat{\beta}, \hat{\delta}_2)$ denotes [Holtz-Eakin et al. \(1988\)](#)'s estimator. Therefore, in order to test that $\eta = \bar{\eta}$, it suffices to consider a test that rejects whenever

$$\frac{Y_{0,2} - \bar{\eta} - x'_{0,2}\hat{\beta}}{\hat{\delta}_2} - (Y_{0,1} - x'_{0,1}\hat{\beta})$$

is in the extreme tails of such empirical distribution. [Ahn et al. \(2013\)](#), for example, discussed how [Holtz-Eakin et al. \(1988\)](#)'s method can be generalized when there are multiple factors. The idea of combining [Holtz-Eakin et al. \(1988\)](#) with [Conley and Taber \(2011\)](#), although straightforward, does not seem to have been considered elsewhere.

We have considered two alternative methods of inference, one based on $T_0 \rightarrow \infty$ asymptotics, and the other based on $J \rightarrow \infty$ asymptotics. In addition to these two methods, we can also entertain the possibility that if both T_0 and J are large, it may be possible to use the panel technique as in [Bai \(2009\)](#) as well.¹³ See, e.g., [Gobillon and Magnac \(2016\)](#). The latter two procedures are based on the presumption that the researcher takes the linear factor structure seriously, so it may be more powerful than the [Andrews \(2003\)](#)'s test. On the other hand, if a researcher views the linear factor model as just a toy model¹⁴ to illustrate the potential problem of difference-in-differences methods, then she would probably be hesitant to discard the synthetic control method, which may be able to accommodate potentially complicated statistical structures that may go beyond the linear factor model.

The three methods that we discussed here as possible alternatives are all theoretically valid under some asymptotics. Asymptotic validity does not necessarily imply that any given method performs

¹³ If one were to assume that $Y_{0,t}(1) = Y_{0,t}(0) + \beta$, the factor model in Equation (3) becomes

$$\begin{aligned} Y_{j,t} &= \alpha_j + \theta_t + \gamma'_j \delta_t + \epsilon_{j,t}, \quad t = 1, \dots, T_0, \\ Y_{j,T_0+1} &= \alpha_j + \theta_{T_0+1} + \gamma'_j \delta_{T_0+1} + \epsilon_{j,T_0+1}, \quad j = 1, \dots, J, \\ Y_{0,T_0+1} &= \beta + \alpha_0 + \theta_{T_0+1} + \gamma'_0 \delta_{T_0+1} + \epsilon_{0,T_0+1}. \end{aligned}$$

Using the pre-treatment data, one can consistently estimate (α_j, γ'_j) ($j = 0, 1, \dots, J$) and (θ_t, δ'_t) as long as $J, T_0 \rightarrow \infty$. Using the control outcome for the period $t = T_0 + 1$ along with (α_j, γ'_j) consistently estimated, one can consistently estimate $(\theta_{T_0+1}, \delta'_{T_0+1})$, which is possible if $J \rightarrow \infty$. Combining (α_0, γ'_0) as well as $(\theta_{T_0+1}, \delta'_{T_0+1})$, one can make an inference of β .

¹⁴ Indeed, [Abadie et al. \(2010\)](#) (Section 2.2) consider some other model (in addition to the factor model) for motivation of the synthetic control.

reasonably for a given finite sample. A serious Monte Carlo comparison of the relative performance of the three alternatives, which is beyond the scope of the current paper, is required to determine a method to be recommended to practitioners.

5. Conclusions

We considered the performance of the permutation test (placebo test) in the context of the synthetic control method. The symmetry assumption, one of the crucial conditions for the validity of the permutation test, may be violated in synthetic control studies. Using Monte Carlo simulations, we show that the size of the permutation tests can be distorted. The results suggest that even with simple DGPs and rather restrictive distributional assumptions of the error term, as long as aggregate shocks are present, the permutation test in its current form is likely to fail and cannot serve as a proper tool for inference with the synthetic control method. Several possible alternatives were discussed. That being said, we should be careful and repeat an anonymous referee's cautious remark that, while our analysis is from a repeated sampling perspective, the synthetic control literature uses permutation tests in the context of design-based inference, and as such, the permutation tests have exact size.

Acknowledgments: Helpful comments by Alberto Abadie, Bruno Ferman and Guido Imbens are greatly appreciated.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association* 105: 493–505.
- Abadie, Alberto, and Javier Gardeazabal. 2003. The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review* 93: 112–32.
- Ahn, Seung C., Young H. Lee, and Peter Schmidt. 2013. Panel Data Models with Multiple Time-Varying Individual Effects. *Journal of Econometrics* 174: 1–14.
- Andrews, Donald W. K. 2003. End-of-Sample Instability Tests. *Econometrica* 71: 1661–94.
- Bai, Jushan. 2009. Panel Data Models With Interactive Fixed Effects. *Econometrica* 77: 1229–79.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics* 119: 249–75.
- Canay, Ivan A., Joseph P. Romano, and Azeem M. Shaikh. 2017. Randomization Tests under an Approximate Symmetry Assumption. *Econometrica* 85: 1013–1030.
- Conley, Timothy G., and Christopher R. Taber. 2011. Inference with "Difference in Differences" with a Small Number of Policy Changes. *Review of Economics and Statistics* 93: 113–25.
- Doudchenko, Nikolay, and Guido W. Imbens. 2016. *Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis*. NBER Working Paper No. 22791; National Bureau of Economic Research, Cambridge, MA, USA.
- Ferman, Bruno, and Cristine Pinto. 2017. *Revisiting the Synthetic Control Estimator*. New York: Mimeo.
- Fisher, Ronald Aylmer. 1949. *The Design of Experiments*, 5th ed. Edinburgh: Oliver and Boyd.
- Gobillon, Laurent, and Thierry Magnac. 2016. Regional Policy Evaluation: Interactive Fixed Effects and Synthetic Controls. *Review of Economics and Statistics* 98: 535–51.
- Hoeffding, Wassily. 1952. The Large-Sample Power of Tests Based on Permutation of Observations. *Annals of Mathematical Statistics* 23: 169–92.
- Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. 1988. Estimating Vector Autoregressions with Panel Data. *Econometrica* 56: 1371–95.

Nolan, John. 2015. *Stable Distributions—Models for Heavy Tailed Data*. Boston: Birkhauser.

Wikipedia Contributors. Stable distribution. *Wikipedia, The Free Encyclopedia*, https://en.wikipedia.org/w/index.php?title=Stable_distribution&oldid=808375411 (accessed on 10 November 2017).



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).