



Working papers in transport, tourism, information technology and microdata analysis

The Power of the Synthetic Control Method

**Kenneth Carling
Yujiao Li
Editor: Hasan Fleyeh**

Nr: 2016:10

Working papers in transport, tourism, information technology and microdata analysis
ISSN: 1650-5581
© Authors

The Power of the Synthetic Control Method

Kenneth Carling[♦] and Yujiao Li

This version: 2017-01-02

ABSTRACT

The synthetic control method (SCM) is a new, popular method developed for the purpose of estimating the effect of an intervention when only one single unit has been exposed. Other similar, unexposed units are combined into a synthetic control unit intended to mimic the evolution in the exposed unit, had it not been subject to exposure. As the inference relies on only a single observational unit, the statistical inferential issue is a challenge. In this paper, we examine the statistical properties of the estimator, study a number of features potentially yielding uncertainty in the estimator, discuss the rationale for statistical inference in relation to SCM, and provide a Web-app for researchers to aid in their decision of whether SCM is powerful for a specific case study. We conclude that SCM is powerful with a limited number of controls in the donor pool and a fairly short pre-intervention time period. This holds as long as the parameter of interest is a parametric specification of the intervention effect, and the duration of post-intervention period is reasonably long, and the fit of the synthetic control unit to the exposed unit in the pre-intervention period is good.

KEY WORDS: Bootstrap; Comparative case study; Counterfactual analysis; Intervention effect; Monte Carlo Simulation; Statistical Inference

1. INTRODUCTION

In social science, it is frequently the case that a single unit has been exposed to some intervention, either as part of a policy in a small-scale testing state or by happenstance. To learn from such cases to form efficient policies in the future, there is a desire to estimate the effect of the intervention. The synthetic control method (SCM) can be applied to complement and facilitate comparative case studies in this regard (Abadie *et al.*, 2010). In comparative case studies, researchers estimate the evolution of aggregate outcomes (such as mortality rates, average income, crime rates, etc.), for a unit affected by a particular occurrence of the event or intervention of interest, and compare it to the evolution of the aggregates obtained for some control group of unaffected units. The idea of SCM is to form a synthetic control unit (hereafter, SC unit) to the unit exposed to the intervention, where the SC unit is a combination of potential control units (i.e. the donor pool). The SC unit shall mimic the evolution of the treated unit had the treated unit not been exposed to the intervention. By comparing the evolution of the variables of interest for the treated unit and the synthetic control unit, the effect of the intervention can be identified.

[♦] Kenneth Carling is a professor in Statistics and Yujiao Li is a PhD-student in Microdata Analysis at the School of Technology and Business Studies, Dalarna University, SE-791 88 Falun, Sweden. *Corresponding author:* Yujiao Li, e-mail: yli@du.se, phone: +46-72-2336356.

Given the important scope of SCM, it is unsurprising that it has been well received in many areas where comparative case studies are common. The work of Abadie *et al.* (2010) was a forerunner to the surge in the application of SCM, as they outlined the method and provided a pedagogical and convincing estimation of the effect on cigarette consumption, resulting from Proposition 99 – a large-scale tobacco control program implemented in California in 1988. They concluded that annual per-capita cigarette sales in California were about 26 packs lower than what it would have been in the absence of Proposition 99. Thereafter, the method has spread rapidly to other fields, such as biology, social science, political economics, environmental studies, and educational studies. To mention a few examples where SCM has been applied: Coffman and Noy (2011) estimated the long-term impact of the hurricane Iniki disaster on tourism development on a Hawaiian island; Ando (2015) estimated the local impact of nuclear power facilities in Japan; Hinrichs (2012) estimated the effect of the banning of affirmative actions on student composition at University of California; Munasib and Rickman (2015) estimated regional economic impact of the shale gas and tight oil boom in Arkansas (and two other states); Pinotti (2015) estimated the impact of organized crime on social development in the two Italian regions, Calabria and Campania; and finally, Abadie *et al.* (2015) estimated the economic impact of the 1990 German reunification on West Germany.

In the examples above and other applications of SCM, the number of treated units is, at the most, a handful and, more commonly, a single one. It comes naturally for a statistician to ask the question: how can an effect-parameter be estimated with meaningful precision, if only a single unit is at hand? Abadie *et al.* (2010), and others, have tested the null hypothesis of no intervention effect by using permutation tests, and, to the surprise of the authors of this paper, frequently rejected the null hypothesis. Nonetheless, Abadie *et al.* (2010) did not recommend using SCM when the pre-intervention match between the treated unit and the SC unit is poor, or the pre-intervention period is short. Overall, they did not provide any compelling reasons as to why it is possible to identify an intervention effect with SCM, nor did they discuss data requirements and other conditions that might shed some light on when the application of SCM is likely to identify the existence of an actual intervention effect.

In this paper, we aim to study the SCM's power to identify an intervention effect, in the interest of providing an inferential framework for the SCM estimator. Our contributions are three-fold. First, we relax an assumption of SCM and examine the SCM estimator's properties, in terms of unbiasedness and consistency. In our view, Abadie *et al.* (2010) introduced an assumption on SCM that appears overly strict. Second, we conduct a simulation study to understand the power of SCM under various data conditions. We find that the three most critical features turn out to be the discrepancy in the outcome variable between the treated unit and the SC unit in the pre-intervention period, the length of post-intervention period, and the size of the donor pool. Other features, such as discrepancy in covariates, correlation of control units, measurement errors, the time-property of the intervention effect, length of pre-intervention period, and model fit between outcome and covariates are less critical. Having said this, it is nevertheless the case that all these features matter and anyone intending to use the SCM for a specific case study is in the dark on whether the method will identify a potential effect or not for the case at hand. To aid in such cases, we have developed a web

application¹, where the case characteristics can be inserted and the case specific power of SCM is returned. As a third contribution of this paper, we extend the SCM point estimator by illustrating on the case study of estimating the effect of the tobacco control program (Abadie *et al.*, 2010) how an interval estimator can be computed and interpreted.

The paper proceeds as follows. Section 2 outlines SCM and its estimation procedure. In Section 3, we extend the original work (Abadie *et al.*, 2010), by relaxing an assumption and derive some properties of the SCM estimator. In Section 4, we study the power of SCM by simulation experiments. This section includes the experimental design, data generation procedure and the simulation results. In Section 5, we re-examine the case study of Abadie *et al.* (2010), by offering interval estimates. Section 6 concludes the paper.

2. THE SYNTHETIC CONTROL METHOD

The parameter of interest is the intervention effect (or treatment effect), which is the difference between the treated unit's outcome of a variable of interest and its counterfactual, had the unit not been treated. To estimate the outcome for the counterfactual, SCM constructs a counterfactual SC unit by weighing control units, such that the weighted average of outcomes and relevant covariates in the pre-intervention period is similar to the factual ones of the treated unit. In outlying SCM in this section, we have tried to adhere to the notations introduced by Abadie *et al.* (2010), but our introduction of some new concepts has forced us to make some alterations.

2.1 Model

The observed outcome variable of interest is denoted by y_{jt} for unit j at time t , where 0 refers to the treated unit and $j = 1, \dots, J$ refers to the J control units in the “donor pool”. Furthermore, $t = 1, \dots, T_0, \dots, T$ is the time period during which the variables are observed, and T_0 is the time point the intervention occurs. η_{0t} denotes the outcome of the variable of interest, had the treated unit not been treated (i.e. the counterfactual). The estimator of $\hat{\eta}_{0t}$ is the outcome for the SC unit, being the weighted average of control units, supposedly mimicking the evolution of the treated unit in absence of treatment;

$$\hat{\eta}_{0t} = \sum_{j=1}^J w_j y_{jt} \quad (1)$$

with the constraints that $\sum_{j=1}^J w_j = 1$ and $0 \leq w_j \leq 1$. $\mathbf{w} = (w_1, w_2, \dots, w_J)'$ is the parameter vector of the units' weights. It is stipulated that the intervention bears effect at the time point of its introduction and afterwards and therefore, for $t < T_0$, it holds that $y_{0t} - \eta_{0t} = 0$. While, for $t \geq T_0$, the parameter of intervention effect equals

$$\alpha_t = y_{0t} - \eta_{0t}. \quad (2)$$

The estimator of the intervention effect at time t is

$$\hat{\alpha}_t = y_{0t} - \hat{\eta}_{0t} = y_{0t} - \sum_{j=1}^J w_j y_{jt}, \quad (3)$$

where w_j is obtained by optimizing an objective function, which minimizes the discrepancy between the observed treated unit and SC unit before the intervention, with regard to the outcome variable and covariates. Abadie *et al.* (2010) also considered covariates to the model,

¹ https://yujiao1026.shinyapps.io/SCM_120116/

where the outcome variable is explained by covariates using a linear factor model. The linear model for control units eq. (4), and the counterfactual of the treated unit eq. (5) are

$$y_{jt} = \delta_t + \boldsymbol{\theta}_t \mathbf{z}_j + \boldsymbol{\lambda}_t \boldsymbol{\mu}_j + \epsilon_{jt}, \quad (4)$$

$$\eta_{0t} = \delta_t + \boldsymbol{\theta}_t \mathbf{z}_0 + \boldsymbol{\lambda}_t \boldsymbol{\mu}_0 + \epsilon_{0t}, \quad (5)$$

where δ_t is an unknown common factor with constant factor loadings across unit, \mathbf{z}_i is a $r \times 1$ vector of r observed covariates, $\boldsymbol{\mu}_i$ is a $F \times 1$ vector of unobserved covariates, and ϵ_{jt} and ϵ_{0t} are error terms with common variance σ^2 .

Note that α_t is allowed to vary freely in the post-intervention period, as Abadie *et al* (2010) suggested estimating it non-parametrically. As will be shown below, such a choice might weaken the power of SCM substantially, and we will therefore also investigate simpler specifications of the intervention effect over time. Box and Tiao (1975) classified the pattern of intervention effects as “step” and “pulse”, where the former implies a constant intervention effect over time. Figure 1 depicts three simple types of the intervention effect after T_0 . These types are useful simplifications, as they can provide a reasonable approximation to the evolution of the intervention effect, while keeping the number of parameters to estimate low, and offer sufficient information for decision-making. In the following, we will focus on the first type being a time-constant intervention effect, and define the parameter α_l which is the averaged post-intervention effect: $\alpha_l = \frac{1}{T-T_0+1} \sum_{t=T_0}^T \alpha_t$. In Section 5, we give an illustration of the third type of intervention effect, namely when the effect of the intervention increases gradually after T_0 .

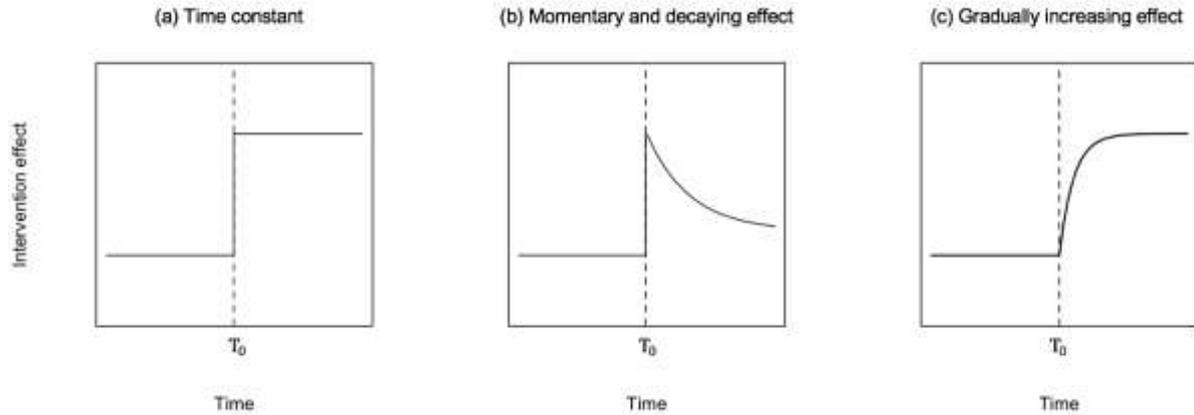


Figure 1: Three types of evolution of the intervention effect

2.2 Estimation method

Neglecting the covariates for the time being, the SCM estimator $\hat{\alpha}_t = y_{0t} - \sum_{j=1}^J w_j y_{jt}$ is a random variable, because it is a function of the random variables y_{jt} as stipulated by the factor model in (4). The weights are estimated by the data in the pre-intervention period and thus stochastically unrelated to $\hat{\alpha}_t$, while y_{0t} is observed and therefore considered non-random. The vector of weights, \mathbf{w} , is obtained by minimizing the discrepancy in the variable of interest between the treated unit and the SC unit in the pre-intervention period. However, in the presence of covariates such a discrepancy is measured in terms of the outcome variable and the covariates simultaneously. Therefore, two objective functions, \mathbf{f}_y and \mathbf{f}_w , are

separately formulated for the outcome variable and the covariates. The first objective function is $\mathbf{f}_v = \|\mathbf{y}_0 - \mathbf{Y}_1 \mathbf{w}\|^2 = (\mathbf{y}_0 - \mathbf{Y}_1 \mathbf{w})' (\mathbf{y}_0 - \mathbf{Y}_1 \mathbf{w})$, where \mathbf{y}_0 $((T_0 - 1) \times 1)$ and \mathbf{Y}_1 $((T_0 - 1) \times J)$ are the outcome variable of the treated unit and the J control units in the pre-intervention period. In detail we have

$$\mathbf{y}_0 = \begin{pmatrix} y_{0,1} \\ y_{0,2} \\ \vdots \\ y_{0,T_0-1} \end{pmatrix} \text{ and } \mathbf{Y}_1 = \begin{pmatrix} y_{11} & y_{21} & \cdots & y_{J1} \\ y_{12} & y_{22} & & \\ \vdots & & \ddots & \\ y_{1,T_0-1} & y_{2,T_0-1} & \cdots & y_{J,T_0-1} \end{pmatrix}.$$

As for the objective function regarding the covariates, it is $\mathbf{f}_w = \|\mathbf{z}_0 - \mathbf{V} \cdot \mathbf{Z}_1 \cdot \mathbf{w}\|^2 = (\mathbf{z}_0 - \mathbf{Z}_1 \cdot \mathbf{w})' \mathbf{V} (\mathbf{z}_0 - \mathbf{Z}_1 \cdot \mathbf{w})$, where \mathbf{z}_0 $(r \times 1)$ and \mathbf{Z}_1 $(r \times J)$ are the r covariates (assumed time-constant, see footnote no. 4) of the treated unit and the control units in the pre-intervention period, and \mathbf{V} is explained below. Moreover, we have

$$\mathbf{z}_0 = \begin{pmatrix} z_{01} \\ z_{02} \\ \vdots \\ z_{0r} \end{pmatrix} \text{ and } \mathbf{Z}_1 = \begin{pmatrix} z_{11} & z_{21} & \cdots & z_{J1} \\ z_{12} & z_{22} & & \\ \vdots & & \ddots & \\ z_{1r} & z_{2r} & \cdots & z_{Jr} \end{pmatrix}.$$

The element z_{jk} is measured by a linear function of covariate $x_{jk,t}$ over the pre-intervention period: $z_{jk} = g(x_{jk,t})$, where $x_{jk,t}$ is the j^{th} unit's k^{th} covariate at time t , and $g(x_{jk,t}) = \sum_{t=1}^{T_0-1} \iota_t \cdot x_{jk,t}$ with the constraints that $\iota_t \in [0,1]$ and $\sum_{t=1}^{T_0-1} \iota_t = 1$. In general, and for simplicity, we take $\iota_t = \frac{1}{T_0-1}$, while in an empirical study ι_t can take on a value of the researcher's choice.

\mathbf{V} is a $r \times r$ symmetric and positive semidefinite matrix, which assigns weights to all the covariates:

$$\mathbf{V} = \begin{pmatrix} v_1 & & & \\ & v_2 & & \\ & & \ddots & \\ & & & v_r \end{pmatrix}.$$

After having given these two objective functions, we will outline the optimization procedure, which is a nested loop of these two functions implemented in the following three steps: First, an initializing matrix for \mathbf{V} is plugged into the function \mathbf{f}_w for optimization, to obtain local optimal value of f_w and for the vector \mathbf{w} . Second, plug the updated \mathbf{w} into function \mathbf{f}_v to obtain a value for f_v . Third, iterate step 1 and 2 until \mathbf{f}_v achieves its minimum. The matrix \mathbf{V} is initialized as $(\mathbf{Z}_1' \mathbf{Z}_1)^{-1} \mathbf{Z}_1' \mathbf{z}_0$ if the matrix $\mathbf{Z}_1' \mathbf{Z}_1$ is reversible, otherwise one uses the $(r \times r)$ identity matrix scaled by $\frac{1}{r}$. The optimization algorithms used are BFGS (Broyden, Fletcher, Goldfarb, and Shanno, 1970), and Nelder-Mead (Nelder and Mead, 1965). The pseudo R-code for the optimization is shown below:

```

fv <- function(v){
  w <- optimize( fw( w, V ) )
  fv.value <- fw( w,V )
  return(w, fv.value)
}

V[1] <- V0
for (i in 2:n){
  if(fv(V[i])$fv.value < fv(V[i-1])$fv.value){
    w.hat <- w[i]
  }else{
    w.hat <- w[i-1]
  }
  update V[i+1]
}

```

3. Some statistical properties of the estimator of the intervention effect

Abadie *et al.* (2010) introduced an overly rigid assumption, unlikely to be met in applied work, and we begin this section by relaxing this assumption of SCM. Thereafter, we show that the estimator of the intervention is unbiased, also, after having relaxed the assumption. We end this section by examining under what conditions the estimator is consistent.

3.1 Relaxing an assumption of SCM

To ensure the unbiasedness of the SCM estimator, Abadie *et al.* (2010) introduced the assumption that the discrepancy of the outcome variable of interest between the treated unit and SC unit was exactly zero in the pre-intervention period. Specifically, $\hat{\alpha}_t = y_{0t} - \sum_{j=1}^J w_j y_{jt}$ ($t \geq T_0$) is asymptotically unbiased under the assumptions that there is a vector \mathbf{w} , such that the following three equations are simultaneously satisfied; $\sum_{j=1}^J w_j y_{jt} = y_{0t}$ ($t < T_0$), $\sum_{j=1}^J w_j \mathbf{z}_j = \mathbf{z}_0$, and $\sum_{j=1}^J w_j \boldsymbol{\mu}_j = \boldsymbol{\mu}_0$.

In practice, however, it is unlikely to find control units in the donor pool that allow for exact matching of the SC unit and the treated unit, with regard to all the three discrepancies, in terms of outcome, observed and unobserved covariates. This issue was discussed by Abadie *et al.* (2010), and they suggested the assumption to be approximately valid, i.e. $\sum_{j=1}^J w_j y_{jt} \approx y_{0t}$, without offering related theoretical work in support of the suggestion. Instead, we suggest that the three conditions are replaced by their counterparts, in terms of expectation, i.e. the conditions are: $E(\sum_{j=1}^J w_j y_{jt}) = y_{0t}$ ($t < T_0$); $E(\sum_{j=1}^J w_j \mathbf{z}_j) = \mathbf{z}_0$ and $E(\sum_{j=1}^J w_j \boldsymbol{\mu}_j) = \boldsymbol{\mu}_0$. Below, we show that this modification preserves the unbiasedness of $\hat{\alpha}_t$ and, consequently, also of $\hat{\alpha}_l$.

3.2 Unbiasedness

We will show that unbiasedness holds in the long-run and at a specific time point, i.e. $E(\hat{\alpha}_l) = \alpha_l$ and $E(\hat{\alpha}_t) = \alpha_t$. Recalling that $\hat{\alpha}_l = \frac{1}{T-T_0+1} \sum_{t=T_0}^T \hat{\alpha}_t$, it follows that $\hat{\alpha}_l$ will be unbiased if $\hat{\alpha}_t$ is unbiased. To show $E(\alpha_t - \hat{\alpha}_t) = 0$ ($t \geq T_0$), we note from eq:s (1,2,4,5) that

$$\alpha_t - \hat{\alpha}_t = \boldsymbol{\theta}_t \mathbf{z}_d + \boldsymbol{\lambda}_t \boldsymbol{\mu}_d + \epsilon_d \quad (7)$$

where \mathbf{z}_d , $\boldsymbol{\mu}_d$ and ϵ_d are respectively the discrepancy between treated unit and SC unit, with regard to observed covariates, unobserved covariates and error terms (i.e. $\mathbf{z}_d = \mathbf{z}_0 - \sum_{j=1}^J \widehat{w}_j \mathbf{z}_j$, $\boldsymbol{\mu}_d = \boldsymbol{\mu}_0 - \sum_{j=1}^J \widehat{w}_j \boldsymbol{\mu}_j$, and $\epsilon_d = \epsilon_{0t} - \sum_{j=1}^J \widehat{w}_j \epsilon_{jt}$). From the assumptions in Section 3.1, we have that $E(\mathbf{z}_d) = 0$, $E(\boldsymbol{\mu}_d) = 0$ and, also, the error terms have expectation equal to zero, i.e. $E(\epsilon_d) = 0$, we have $E(\widehat{\alpha}_t) = \alpha_t$ ($t \geq T_0$), and it follows $E(\widehat{\alpha}_l) = \alpha_l$.

3.3 Consistency

In a practical application in which SCM would be used, the researcher has observations on y_{0t} and a set of possible control units. In principle, this means that the length of the post-intervention time ($\Delta T = T - T_0 + 1$) could tend to infinity, if the researcher is sufficiently patient. It also means that, in principle, the size of the donor pool could be arbitrarily large, i.e. J could tend to infinity. Though the SCM estimator is unbiased, the precision in this estimator of an intervention effect with regard to ΔT and J is unclear. Thus far, no one, to the best of our knowledge, has tried to investigate the precision of the SCM estimator. In making such investigation, we will first consider the consistency of the estimator and complement this analysis with subsequent simulations for finite donor pools and post-intervention length.

An additional reason for this analysis is that the criteria for choosing J and ΔT are lacking in most applications of SCM. Abadie *et al.* (2015) refined the donor pool by discarding control units with comparatively poor fit in pre-intervention period, but the appropriate size of a donor pool has never been discussed. We aim to assist researchers in the choice of J and ΔT , an issue we come back to in the subsequent section.

As the SCM estimator is unbiased, this analysis will focus on the variance of the estimator, and we start by studying the precision as the post-intervention length tends to infinity. The variance of $\widehat{\alpha}_t$ and $\widehat{\alpha}_l$ are derived as follows: from $\widehat{\alpha}_t = y_{0t} - \widehat{\eta}_{0t} = y_{0t} - \sum_{j=1}^J w_j y_{jt}$, we have that $\text{var}(\widehat{\alpha}_t) = \text{var}[y_{0t} - \sum_{j=1}^J w_j (\delta_t + \boldsymbol{\theta}_t \mathbf{z}_j + \boldsymbol{\lambda}_t \boldsymbol{\mu}_j + \epsilon_{jt})] = \text{var}(\sum_{j=1}^J w_j \epsilon_{jt})$. Now $\text{var}(\epsilon_{jt}) = \sigma^2$ and, thus, $\text{var}(\widehat{\alpha}_t) = \sigma^2 \sum_{j=1}^J w_j^2$ and $\text{var}(\widehat{\alpha}_l) = \frac{1}{\Delta T^2} \sum_{t=T_0}^T \text{var}(\widehat{\alpha}_t) = \frac{\sigma^2 \sum_{j=1}^J w_j^2}{\Delta T}$, since the error terms, ϵ_{jt} , are assumed independent (cf eq. (4)).

As ΔT goes to infinity, $\widehat{\alpha}_l$ converges to α_l , whereas $\widehat{\alpha}_t$ does not converge to the corresponding parameter. To show the consistency of $\widehat{\alpha}_l$, we use the squeeze theorem that states: *if $a_n \leq b_n \leq c_n$ for sufficiently large n and $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = x$, then $\lim_{n \rightarrow \infty} b_n = x$* . Hence, if $0 \leq \text{var}(\widehat{\alpha}_l) \leq \frac{\sigma^2}{\Delta T}$ and $\lim_{\Delta T \rightarrow \infty} \frac{\sigma^2}{\Delta T} = 0$, then it follows that $\lim_{\Delta T \rightarrow \infty} \text{var}(\widehat{\alpha}_l) = 0$, which implies that $\widehat{\alpha}_l$ is consistent. Therefore, it needs to be proven that $\text{var}(\widehat{\alpha}_l) \leq \frac{\sigma^2}{\Delta T}$. To make this proof, note first that $w_j^2 \leq w_j$, as $w_j \in [0,1]$, which in turn implies that $\sum_{j=1}^J w_j^2 \leq \sum_{j=1}^J w_j = 1$ and therefore $\text{var}(\widehat{\alpha}_l) = \frac{\sigma^2 \sum_{j=1}^J \widehat{w}_j^2}{\Delta T} \leq \frac{\sigma^2}{\Delta T}$.

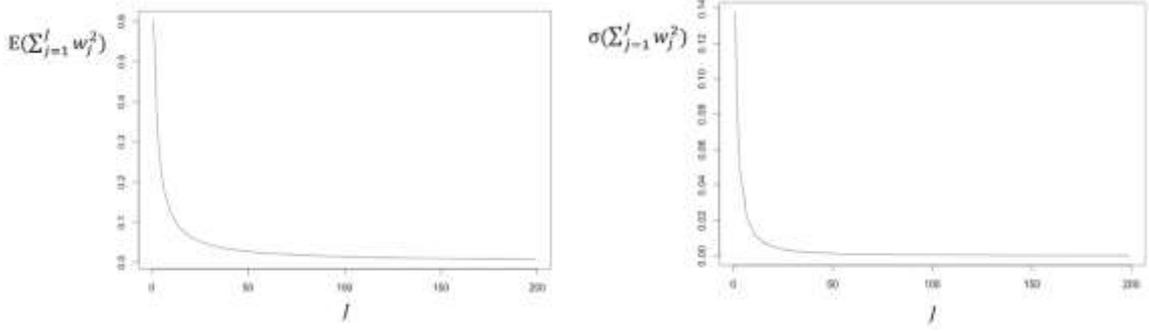


Figure 2: Rate of convergence of $\sum_{j=1}^J w_j^2$ as a function of J in the numerical example.

A sufficient condition for both $\hat{\alpha}_l$ and $\hat{\alpha}_t$ to be consistent as J tends to infinity is that $\lim_{J \rightarrow \infty} \text{var}(\hat{\alpha}_l) = 0$, and $\lim_{J \rightarrow \infty} \text{var}(\hat{\alpha}_t) = 0$. However, $\text{var}(\hat{\alpha}_t) = \sigma^2 \sum_{j=1}^J \hat{w}_j^2$ and $\text{var}(\hat{\alpha}_l) = \frac{\sigma^2 \sum_{j=1}^J \hat{w}_j^2}{\Delta T}$ will tend to zero, if $\lim_{J \rightarrow \infty} \sum_{j=1}^J \hat{w}_j^2 = 0$. In the special case of equal weights $w_i = w_j$ ($\forall i, j$), it follows $w_j = \frac{1}{J}$ and thus $\sum_{j=1}^J w_j^2 = \frac{1}{J}$, with the implication that $\lim_{J \rightarrow \infty} \sum_{j=1}^J w_j^2 = 0$. Intrinsic in this argument is that the donor pool is infinite, and there are infinitely many control units with an associated positive weight. For the case of unequal weights of the control units, a demonstration of consistency is messy, and instead we provide a numerical example that also gives an understanding of the rate of convergence. Figure 2 gives the mean and the standard deviation of $\sum_{j=1}^J w_j^2$, as a function of J from the numerical example. The numerical example was generated accordingly: J standard uniform random variates were generated, and thereafter scaled by their sum to meet the restriction that $\sum_{i=1}^J w_i = 1$. Starting at one, J was increased in steps of one up to 200, and for each J the procedure was replicated 200 times.

4. POWER STUDY AND SIMULATION

Even though the SCM estimators are unbiased, and in some instances consistent², it does not mean that SCM is a powerful approach for a particular application at hand. In this section, we study the method's power under various scenarios of data features, and investigate which features mainly affect the power. We focus on the variance of the SCM estimator and apply Monte Carlo simulations in the investigation. The reason for resorting to simulations is that neither $\text{var}(\hat{\alpha}_t)$ nor $\text{var}(\hat{\alpha}_l)$ can be analytically expressed in finite sample scenarios, because their variance depends on w_j , which is retrieved from a heuristic search and, thus, not amenable to analytical calculations.

The structure of this section is that, first, the data generating process is described and, second, some typical simulation results are presented. As an exhaustive simulation covering all relevant scenarios in which the SCM might be applied would generate massive amounts of

² We have conducted simulation experiments to confirm the unbiasedness and consistency properties, when the SCM implementation is used, since it builds on numerical procedures with hard to foresee performance.

results, we have limited the presentations to a few scenarios, and refer instead to the web app that we have developed for researchers to pre-test the power of SCM for their data.³

In the data -generating process and in the simulations, we have varied five data features which potentially influence the variance of the SCM estimator. The features (with the corresponding control parameters) in the simulation are:

- similarity between the treated and the SC unit in the outcome variable (ρ_y)
- similarity between the treated and the SC unit in covariates (ρ_x)
- correlation in the outcome variable between the control units (ω)
- time dependence in the outcome variable and the covariates (τ)
- explanatory power of the covariates (ρ_{xy})

Furthermore, the analytical investigation of the statistical properties of the SCM estimator suggests that the length of the post-intervention period (ΔT), the size of the donor pool (J), and the distribution of weights (w_j) are additional, important features, whereas the length of the pre-intervention period presumably is inconsequential, but nonetheless considered in the simulations.

The control parameters relate to typical challenges arising in applied work, as measurement errors in the variables, the difficulty of defining the donor pool and obtaining data for relevant control units, the fitting of the g -function, and the cross-over between units. For instance, large measurement errors in the covariates and the outcome variable would decrease the explanatory power of the covariates, and such cases could be investigated by lowering the control parameter ρ_{xy} .

4.1 Data Generating Process

The specific details of the data generating process are as follows. We use two covariates for the control units $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, where

$$\mathbf{X}^{(1)} = \left(\mathbf{x}_1^{(1)}, \dots, \mathbf{x}_J^{(1)} \right) = \begin{pmatrix} x_{11}^{(1)} & x_{21}^{(1)} & x_{J1}^{(1)} \\ x_{12}^{(1)} & x_{22}^{(1)} & x_{J2}^{(1)} \\ \vdots & \vdots & \vdots \\ x_{1T}^{(1)} & x_{2T}^{(1)} & \dots & x_{JT}^{(1)} \end{pmatrix}.$$

$x_{jt}^{(1)}$ is the first covariate of the j^{th} control unit at time t . $\mathbf{x}_j^{(1)}$ is generated as an AR(1) process as $x_{jt}^{(1)} = \tau \cdot x_{j,t-1}^{(1)} + \varepsilon_{jt}^{(1)}$, with the error term normal with expectation equal to zero and variance equal to $(1 - \tau^2)$. To introduce correlation between units, we generate $\mathbf{x}_j^{(1)}$ by adding a time series vector $\mathbf{x}_c^{(1)}$, such that $\mathbf{x}_j^{(1)} = \sqrt{\omega} \mathbf{x}_c^{(1)} + \sqrt{(1 - \omega)} \mathbf{x}_j^{(1)}$, where $\mathbf{x}_c^{(1)} \sim \text{AR}(1)$. The same goes for $\mathbf{X}^{(2)}$. The covariates of the treated unit are the linear combination of all control units, $\mathbf{x}_0^{(m)} = \sqrt{\rho_x} \mathbf{X}^{(m)} \mathbf{w} + \sqrt{(1 - \rho_x)} \cdot \boldsymbol{\varepsilon}$, $m = 1, 2$. \mathbf{w} can be set

³ https://yujiao1026.shinyapps.io/SCM_120116/

as either a vector of equal weights, $\mathbf{w} = \frac{1}{J}(1, 1, \dots, 1)'$, or a vector of random weights, (where the randomization is identical to the procedure in the numerical example in subsection 3.3).⁴

Moreover, y_{jt} is generated by a linear model: $y_{jt} = \sqrt{\rho_{xy}} [\beta_1 x_{jt}^{(1)} + \beta_2 x_{jt}^{(2)}] + \sqrt{(1 - \rho_{xy})} \cdot \varepsilon$, where again the error term is standard normal. The counterfactual of the treated unit is generated as $\eta_{0t} = \sqrt{\rho_y} \sum_{j=1}^J w_j y_{jt} + \sqrt{(1 - \rho_y)} \cdot \varepsilon_t$, while the factual outcome in the treated unit is $y_{0t} = \eta_{0t} + \alpha$, with α set to zero in the pre-intervention time period, and to a positive constant in the post-intervention period. All the control parameters are in the range from zero to one.

4.2 Simulation and results

We begin by illustrating the web app for examining the power of SCM on a case with certain features. Figure 3 shows the interface of the web app. In the example, the features are that the donor pool consists of 30 control units, all with positive and equal weights, and the outcome variable has been measured yearly for 30 years, of which 20 are prior to the intervention and 10 are after the intervention, and the true time-constant intervention effect amounts to 2 (i.e. $J = 30, T = 30, T_0 = 20$ and $\alpha = 2$). Furthermore, there are two covariates with parameters $\beta_1 = \beta_2 = 0.5$ and $\tau = 0.85$, $\omega = 0.65$, $\rho_x = \rho_y = \rho_{xy} = 0.5$, where the last settings implies that the outcome variable has strong time dependence, and that there is a fairly strong correlation in the outcome variable between units.

The scenario's setting in the web app allows for the features $\rho_x, \rho_y, \rho_{xy}, \omega$ and τ to be set to any value in the range 0 to 1, and it also provides two options for setting weights. As an output of the web app, the sampling distribution of the estimators $\hat{\alpha}_{T_0}$ and $\hat{\alpha}_l$ is displayed in order to provide assessment of whether the SCM would have sufficient power to identify an intervention effect, as well as an interval estimator of the effect that exploits bootstrapping.

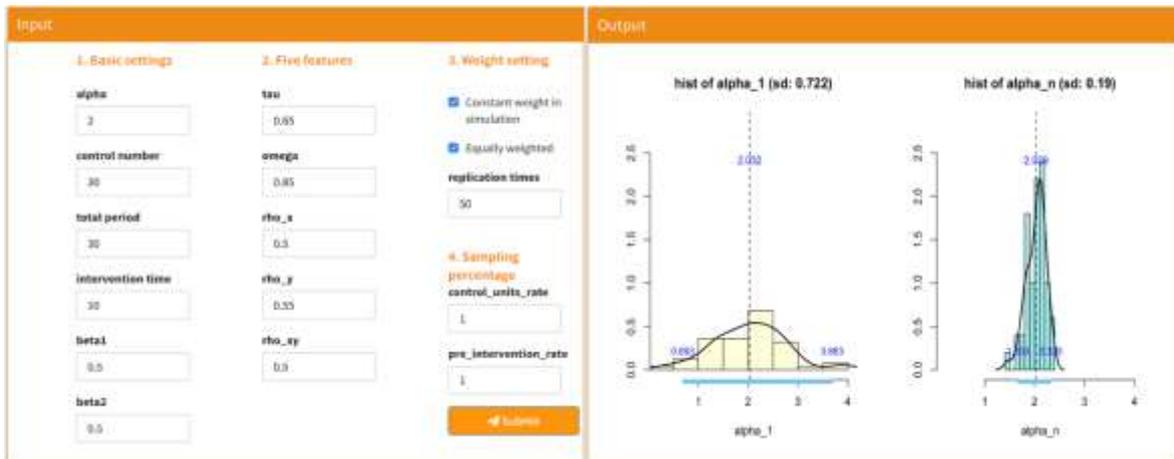


Figure 3: Web-app interface. The left panel illustrates the user's interface for assigning values to the features, whereas the right panel gives the simulated sampling distributions of $\hat{\alpha}_{T_0}$ and $\hat{\alpha}_l$.

⁴ Note that at present, the software implementation of SCM in R and STATA does not allow for time-varying covariates, as all covariates are averaged over the pre-intervention period.

Furthermore, we have conducted a simulation experiment to be able to provide some general guidelines, with regard to crucial features for the power of SCM. Again, we have taken $J = 30, T = 30, T_0 = 20$, as we already know from the analysis of the SCM estimator's statistical properties that increases in the donor pool and the post-intervention length will increase the power. We have again set $\alpha = 2$. The focus is on how the other features affect the power, and we have therefore varied the features $\rho_x, \rho_y, \rho_{xy}, \omega$, and τ from zero to one, one feature at the time keeping the other four fixed⁵.

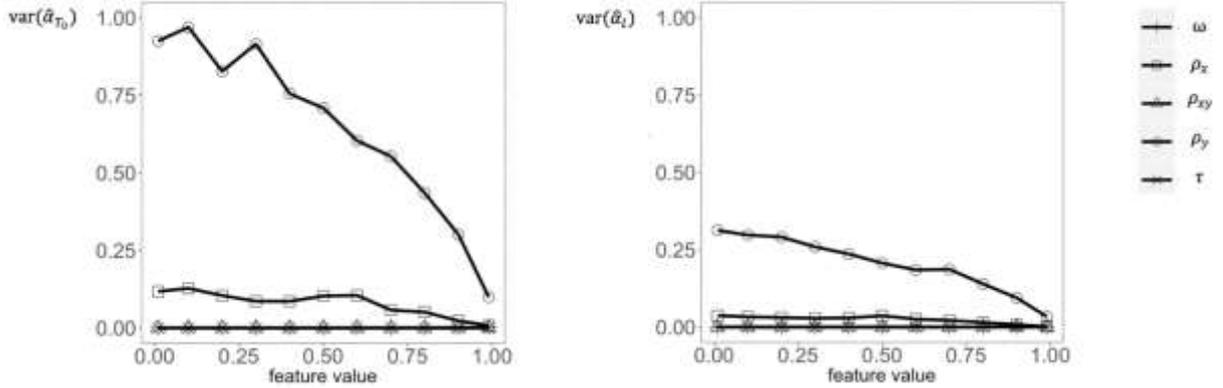


Figure 4: Variance of the estimator as a function of the five features. The left panel shows the variance for the non-parametric estimator (at T_0), whereas the right panel shows the variance for an estimator a time-constant intervention effect.

Figure 4 gives the variance of the non-parametric estimator of the time point, after the intervention in the left panel, as well as the variance of \hat{a}_t as a function of the five features. The first thing to note is the drastically smaller variance of the latter estimator. The second thing to note is that only ρ_y and ρ_x seem to influence the variance of the SCM estimator, and mostly so ρ_y . For this reason, we conclude that the discrepancy between the treated and the SC unit in the outcome variable is the greatest source for uncertainty in the estimate of the intervention effect. Of course, this analysis presumes that the features are independent, which may not be the case and, therefore, we proceed to do a factorial experiment. In the factorial experiment, we vary eight features with the settings as follows: $J = 10, 30; T_0 = 10, 40; \Delta T = 1, 5, 20; \rho_x = 0.1, 0.5; \rho_y = 0.1, 0.5, 0.9, \rho_{xy} = 0.1, 0.5; \omega = 0, 0.5; \tau = 0, 0.5$, while the weights are randomly assigned.

The simulation experiment was conducted for all 576 combinations, with 200 replications for each combination. The features related to the number of control units, as well as the pre- and post-intervention time length were transformed such as to render all features bounded by the interval zero to one before an ANOVA. Table 1 gives the estimates of a regression of the logarithm of $\sigma(\hat{a}_t)$ per experimental combination on the features, as the preceding ANOVA indicated weak interaction effects. Although the features examined generally exercised some influence on the precision of the estimate of the post-intervention effect, three features stand

⁵ The fixed values of the features control parameters are: $\rho_x = 1, \rho_y = 1, \rho_{xy} = 1, \omega = 0$, and $\tau = 0$.

out: ρ_y (discrepancy between the treated and the SC unit in the outcome variable), and ΔT (length of the post-intervention period), and, to some extent, the size of the donor pool (J).

Table 1: Regression model of the logarithm of the standard deviation for $\hat{\alpha}_l$ on the features. Data comes from the factorial experiment. Standard errors in parentheses.

Feature	All features	Key features
$1/\sqrt{\Delta T}$	1.83 (0.02)	1.83 (0.02)
$1/\sqrt{T_0}$	0.10 (0.09)	
$1/\sqrt{J}$	0.52 (0.10)	0.52 (0.10)
ρ_x	-0.11 (0.03)	
ρ_y	-1.13 (0.02)	-1.13 (0.02)
ρ_{xy}	-0.03 (0.03)	
ω	0.20 (0.03)	
τ	-0.03 (0.02)	
R^2	0.949	0.943

5. RE-EXAMINING THE INTERVENTION EFFECT OF CALIFORNIA'S PROPOSTION 99

As stated in the introduction, we were surprised to find so many applications of SCM that, based on permutation tests, rejected the null hypothesis of no intervention effect. Abadie's *et al.* (2010) study of the effect of California's tobacco control program was such an example. In this section we re-analyze this data to provide an interval estimator for the intervention effect, and examine the features to understand why the intervention effect was found statistically significant.

California's proposition 99 was an anti-tobacco legislation, which increased the excise tax on cigarettes, with the tax revenues to be used for tobacco mitigation measures. The data used by Abadie *et al.* (2010) was annual state-level panel data for the period 1970–2000, for 39 US states. Proposition 99 was passed in November 1988, and went into effect in January 1989, which implies 19 years of pre-intervention data (i.e. $T_0 = 19$). The outcome variable was per capita cigarette sales in packs, covariates variables were average retail price of cigarettes, per capita state personal income (in logs), the percentage of the population age from 15 to 24, and per capita beer consumption. The estimated intervention effect was about 26 annual averaged per-capita cigarette sales packs reduction (corresponding to a reduction of about 32%). The estimate of $\hat{\alpha}_l = -26$ was obtained with the donor pool size of 38. We use their data and replicate their procedure to get the intervention effect estimate, as shown in Figure 5 (the left panel is cigarette sales in California and the synthetic California). The right panel shows the discrepancy in cigarette sales between California and the synthetic California, where the estimate refers to the effect of an absence of Proposition 99.

The value of the features' parameters in the dataset are $J = 38$; $T_0 = 19$; $\Delta T = 12$; $\rho_y = 0.99$; $\rho_{xy} = 0.63$, and ρ_x of the covariates are respectively 0.948 (ln GDP per capita), 0.996 (retail price), 0.719 (beer consumption per capita), and 0.998 for the percentage aged 15-24, after having standardized the variables. ω and τ are not computed, due to the partially missing

data in the covariates. The observation that the control parameters are close to one, and ΔT and J are not small, suggests that SCM should be powerful in detecting large intervention effects. More specifically, the explanation as to why SCM proved powerful in this case lies in the fact that ρ_Y was close to one, and the intervention effect was substantial. In fact, by examining the trend after 1975 in cigarette sales in California, and the other states that were assigned a positive weight in the original analysis, one sees a common, almost perfectly linear, down-trend in sales. To detect a break in the linear trend in cigarette sales is not difficult, and this is the crude explanation as to why the SCM was powerful in this intervention effect analysis.

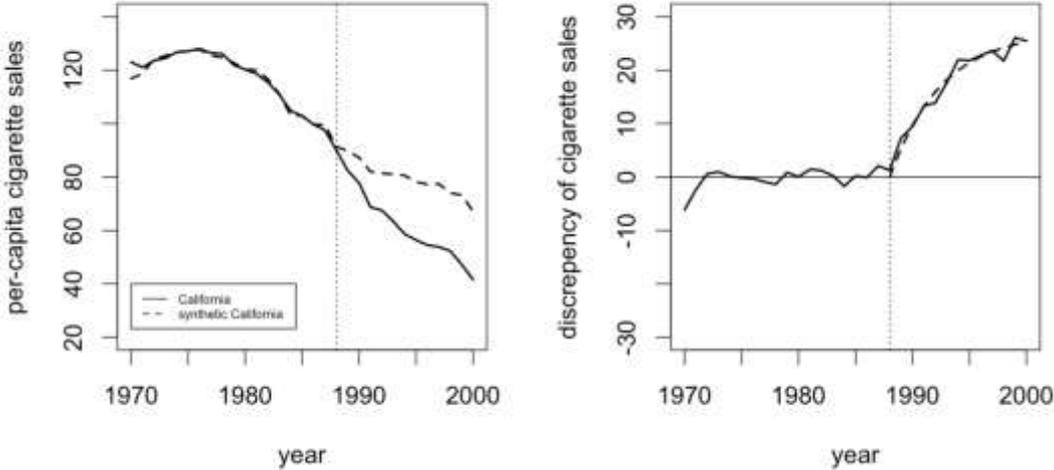


Figure 5: SCM re-estimation of the Proposition 99 absence effect

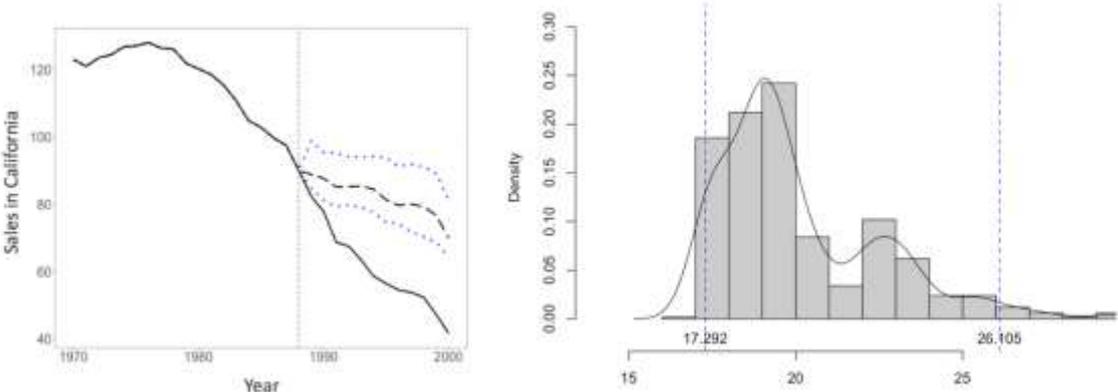


Figure 6: Bootstrap-intervals for the effect of absence of Proposition 99. The dotted line in the left panel is the bootstrap interval for each year, while the right panel is the bootstrap sampling distribution of $\hat{\alpha}_t$.

Next, we use bootstrapping to illustrate an interval estimate of the intervention effect. However, before doing so, it should be recalled that the consistency analysis in Section 3 considered letting ΔT and J tend to infinity, whereas y_{0t} is regarded as observed and thus non-random. The concept of replication is therefore only meaningful, with regard to post-intervention time length and donor pool size, and, for this reason, the inference relates to the effect of not intervening, i.e. the outcome absent Proposition 99. In the bootstrapping procedure, we created 500 bootstrap-samples, by sampling from the 38 states with

replacement, and applied the SCM to obtain an intervention effect estimate for each of the 500 bootstrap-samples. From the resulting distribution of $\hat{\alpha}_l$, the 0.025 and 0.975 quantile served as the interval estimate. Figure 6 shows that interval estimate of annual averaged intervention effect, $\hat{\alpha}_l$, is [17.29, 26.10]. Hence, it can be concluded that without Proposition 99 cigarette sales in California would have been higher.

However, by looking at the non-parametrically estimated effect and recalling the design of the Proposition 99 it is clear that the assumption of a time-constant effect is overly simplistic. We expect an intervention like the Proposition 99 to have a fast effect on cigarette sales as a consequence of the price increase due to the excise tax, while the tobacco mitigating measures funded by the tax revenues would bear effect in a slower manner. We have therefore re-applied the SCM assuming a gradually increasing effect (cf Figure 1c). There are many specifications available (e.g. Xu, 2006), but we choose a simple, two-parameter specification being $\alpha_t = \alpha_0(1 - \exp(-\gamma(t - T_0)))$ where α_0 is the asymptotical effect and γ determines the speed of the intervention to reach its full effect. In the right panel of Figure 5, the fitted effect is imposed as a dashed line. Re-applying the Bootstrap procedure we obtained the estimates 29.58 and 0.23 with the intervals [22.37, 41.48] and [0.13, 0.37] for the parameters α_0 and γ , respectively. The point-estimate of the Proposition 99 is therefore larger than reported by Abadie et al (2010), but clearly subject to much uncertainty.

6. CONCLUSION

The Synthetic control method is a new approach which is rapidly spreading into many academic areas, with the scope of estimating the intervention effect in comparative case studies. In general, it should be expected that the SCM has little power in detecting an intervention effect, and, for this reason, we provide a web app that can be used in advance of planning a study and collecting data, to determine the odds for the SCM to detect an intervention effect for a specific case.

As SCM is typically based on a single unit, subject to intervention, the issue of statistical inference is a concern. This study suggests that statistical inference is meaningful, as long as one realizes that randomization and replication are with regard to the donor pool and the post-intervention time length. For this reason, the interpretation should be the effect of *not* having been exposed to the intervention.

Since SCM provides unbiased estimates, the power of the method is a function of the estimator's variance. This variance is foremost decreasing in ρ_y (i.e. the similarity between the treated and the SC unit in the outcome variable), and with the size of the donor pool. Further, substantial reduction in the variance is readily obtained by replacing the non-parametric estimator proposed by Abadie *et al*, (2010) with simple parametric estimators. Such simple parametric estimators are consistent, and are comparable to the non-parametric estimator when it comes to decision making.

Finally, our re-examination of the Proposition 99 case indicate that the long-term effect was underestimated by Abadie *et al*, (2010), due to an insufficiently long post-intervention period, coupled with their use of a non-parametric estimation approach.

Acknowledgments

The authors would like to thank Dalarna University's internal grant and The Swedish Retail and Wholesale Council who financed this study. The funding source had no involvement in study design, data collection and analysis, or the decision to submit the article for publication. We are also grateful to Moudud Alam, who helped in developing this work and to Niklas Rudholm, Daniel Wikström, and Kristin Svensson for comments both on the work and earlier drafts of this paper.

References

- Abadie, A., Diamond, A. and Hainmueller, J., (2010), Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program, *Journal of the American Statistical Association*, 105, 493-505
- Abadie, A., Diamond, A. and Hainmueller, J., (2015), Comparative Politics and the Synthetic Control Method, *American Journal of Political Science*, 59, 495-510
- Ando, M., (2015), Dreams of urbanization: Quantitative case studies on the local impacts of nuclear power facilities using the synthetic control method, *Journal of Urban Economics*, 85, 68–85
- Box, G. E. and Tiao, G. C. (1975), Intervention analysis with applications to economic and environmental problems, *Journal of the American Statistical association*, 70, 70-79
- Broyden, C.G., (1970), The convergence of a class of double-rank minimization algorithms, *Journal of the Institute of Mathematics and Its Applications*, 6, 76–90
- Coffman, M. and Noy, I., (2011), Hurricane Iniki: measuring the long-term economic impact of a natural disaster using synthetic control, *Environment and Development Economics*, 17, 187–205
- Fletcher, R., (1970), A New Approach to Variable Metric Algorithms, *The Computer Journal*, 13, 317–322
- Goldfarb, D., (1970), A Family of Variable Metric Updates Derived by Variational Means, *Mathematics of Computation*, 24, 23–26
- Hinrichs, P., (2012), The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities, *Review of Economics and Statistics* 94, 712-722

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

Pinotti, P., (2015), The economic costs of organised crime: evidence from southern Italy, *The Economic Journal*, 125, F203–F232

Shanno, D.F., (1970), Conditioning of quasi-Newton methods for function minimization, *Mathematics of Computation*, 24, 647–656

Xu, X. (2006), Modelling and interpreting disease progress in time, In *The Epidemiology of Plant Diseases*, 215-238, Springer Netherlands