國立臺灣大學社會科學院經濟學系

碩士論文

Department of Economics

College of Social Sciences

National Taiwan University

Master Thesis

以蒙地卡羅法探討人工對照組之適用性

Artificial Control Methods: A Monte Carlo Study

陳之翰

Chih-Han Chen

指導教授：陳釗而 博士

Advisor: Jau-Er Chen, Ph.D.

中華民國 102 年 6 月

June, 2013

摘要

本研究透過蒙地卡羅法之模擬以評量 Hsiao et al. (2012) 提出的縱橫資料反事實法(Panel-Data Counterfactual Method)與 Abadie et al. (2010) 提出的合成對照組法(Synthetic Control Method)之人工對照組估計之準確度。在使用主體之選擇由資料產生，與具單位處理效果(Treatment effect)之估計方法中，何者對隨時間變動之處理效果估計較佳為本研究之目標。將此二法，應用於多個實例資料，進行模擬及交叉比對，以評估其可用性。結果顯示當共同因素在時間中的變動與因素負荷在地區間的變動均小時，此二法之估計均良好。但以均方差衡量對人工對照的估計時，縱橫資料反事實法在大多數情況下, 比合成對照組法準確。儘管此二法均需謹慎使用，由蒙地卡羅模擬之結果顯示，縱橫資料反事實法明顯較佳。

關鍵詞：人工對照組估計，模式模擬，縱橫資料反事實法，合成對照組法，蒙地卡羅法

Abstract

The accuracy of the artificial control estimation using the panel-data counterfactual method proposed by Hsiao et al. (2012) and the synthetic control method proposed by Abadie et al. (2010) were evaluated using the Monte Carlo simulations. The aim was to determine which of the methods is superior in studies with time-variant treatment effect, individual treatment effect and data-driven subject selection process. A cross checking process and simulations conducted under various model settings provide guidance on the applicability of these two methods. Both methods perform satisfactory when the variation of common factors in time and factor-loadings across regions are small. In most cases the panel-data counterfactual method is more accurate in artificial control estimation in term of mean-square-deviation criteria than the synthetic control method. Though both methods must be used with caution, the panel-data counterfactual method is clearly the better method suggested by the Monte Carlo results.

Keywords: Artificial control estimation, model simulation, panel-data counterfactual method, synthetic control method, Monte Carlo method

Content

List of Figures

List of Tables

List of Tables, Continued

1. Introduction

Many studies in economics involve determining the effects of a certain event, incident or treatment; for example the study on Taiwan's membership in the WTO is very likely involved with the study of the effect of becoming a WTO member on Taiwan's GDP growth. A review on treatment effect methods by Imbens and Wooldridge (2009) discusses the assumptions and advantages of available methods in the literature. As the pool of the methods grow continuously, it can be difficult for researchers to choose between various methods available. Therefore, comparisons among or between the methods would provide insight and are valuable to researchers. Some of the new methods focus on treatment effect problems in macroeconomic setting and provide estimates of time-variant treatment effect in a targeted region. As these methods are potentially useful in many macroeconomics studies, it is important that their applicability are known to researchers.

An important object of interest in policy evaluation is that the change in certain outcomes of the regions (or subjects, units, individuals and etc.) is affected by some treatments of interest (or incident, laws, policy changes and etc.). To study these treatment effects, two observations of the same unit that are exposed (targeted) or not exposed (control) to the treatment, is required to be known. However, a "fundamental problem of causal inference" (Holland, 1986) is that a unit can only be observed when it

is exposed to a single level of treatment at a time. Data available for evaluations of the treatment effects are therefore limited to different units or the same unit in different times. This problem raises different issues in macro- and microeconomic studies due to the difference in the goal of researches. Studies in macroeconomics tended to target the treatment effect in a certain region and at times required estimates of the trend of the treatment effects in time. These goals could not be accomplished without a control group of the targeted region. Since countries or regions possess unique characteristics that are unlikely to be matched directly by other countries or regions, the lack of control group is an important question in macroeconomic treatment effect studies. Studies in microeconomics on the other hand encountered a different issue as their goals usually were set to find the treatment effects amongst a certain group of individuals, of them some are and some are not affected by the treatment. Taking just the difference in the average of the individuals with varying actions is not appropriate because of the difference in the characteristics of the groups. For example, people who choose to buy health insurances are of different characteristics than those who choose not to. Such difference, if not adjusted, may lead to an inconsistent estimation of the treatment effect.

Two artificial control methods, the SCM and the PDCM, are of interest because they yield the time-variant treatment effect and data-driven selection of control regions. The SCM proposed by Abadie et al. (2010) and the PDCM proposed by Hsiao et al.

(2012) both could generate a weighted sum of the control regions using pre-treatment-period data to fit the pre-treatment-period behavior of the targeted region. Created by combining the generated weighted sum with post-treatment-period data of the control regions, the artificial targeted region (or the counterfactual) mimics the behavior of the targeted region supposing that treatment does not occur. The treatment effects are therefore derived by differentiating the real targeted region with the counterfactual.

Although the SCM and PDCM both produce the counterfactual, their underlying method and model specification are drastically different. The SCM is based on the assumption that the time-variant factor-loadings are the same across regions and the common factors are time-invariants that differ across regions. The PDCM whereas assumes region-specific factor-loading and time-variant common factors. Their minimizing functions in generating the weight are different as well. Though both SCM and PDCM can be used on any data with DID settings, the differences in methodology suggest that they are most likely different in applicability and goodness of post-treatment-period fit.

In this article, the applicability of the methods in different model setting and the differences in goodness of post-treatment-period fit of the two methods were investigated. By utilizing the Monte Carlo Method proposed by Metropolis and Ulam

(1949), the SCM and PDCM were applied to different model structures and variable settings that were generated repeatedly using Matlab 2010a (MathWorks, MA, USA); the mean square deviation (MSD) of the counterfactual and post-treatment-period data of the targeted region were recorded along with the random variables. Regression analysis of the MSD showed that both PDCM and SCM fit better when common factor variance and factor-loading variance are small. In addition, results showed that the PDCM exhibited a mean MSD smaller than did SCM across most model settings. Therefore the PDCM is recommended over the SCM.

The rest of the article is arranged as follows. Section 2 discusses the previous studies on treatment effects. Section 3 discusses the SCM and the PDCMs. Section 4 discusses the generation and analysis of simulation data. The results and discussions are presented in Section 5. Section 6 concludes.

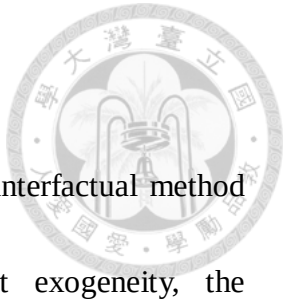
2. Literature Review

The study of the treatment effect was pioneered by Ashenfelter (1978) with subsequent improvements by Ashenfelter and Card (1985), Heckman and Robb (1985), Lalonde (1986), Fraker and Maynard (1987), Card and Sullivan (1988), and Manski (1990). These studies have investigated mainly labor market programs using observational data and have been focused on the self-selection bias, in which the endogenous differences between individuals were motivated. Traditional methods such as the fix effect method and the instrumental variable method have been implemented to deal with the biases. A subsequent econometrics approach, the semi-parametric models, allows for "fewer functional form and homogeneity assumptions" (Imbens and Wooldridge, 2009). However, there would be no general approach without the assumptions of exogeneity, which is defined by the differences in the observed common factors account for "all biases in comparison between targeted and control unit" (Imbens and Wooldridge, 2009). Without a general method for treatment effect estimation, researchers must develop their own specific methods for different types of datasets. As most of the previous methods have been shown to depend heavily on the average treatment effect on groups for individuals, they are insufficient in estimating accurately the treatment effect on a single targeted region, as often the case in most macroeconomics studies. The correction for self-selection bias is adequate for

estimation of data that the targeted unit and the control unit are both groups with numerous individuals such that the average characteristic except self-selection can be matched after taking the means, whereas in many of the macroeconomic cases, the targeted region is consisted of a single unit with unique characteristics that cannot be matched by any other region alone.

Of the methods proposed for various special cases, the difference-in-difference (DID) method has been widely used in empirical economics with influential applications such as Card (1989) and Card and Krueger (1994). The DID method relies on the setting that the targeted region is affected by the treatment after a certain time period (treatment period), whereas the control region is never affected. The treatment effects are obtained by differentiating the growth in the targeted region to the growth in the control region. This double difference eliminates the bias of the permanent differences between the targeted and control regions and the bias originated from time differences. The selection of the control region is, however, often ambiguous and subjective. The credibility of researches would increase if the selection processes are objective and clear, and the researchers could be relieved from repeated test using different control region groups. Moreover, the effects of many treatments are time-variant. For example, the drop in cigarette consumption due to an increase in tobacco taxes is expected to increase over time, because smokers need time to quit

smoking. Therefore, methods that could estimate effectively time-variant treatment effect are valuable to those studies.



3. Methods

Both the synthetic control method (SCM) and the panel-data counterfactual method (PDCM) deal with macroeconomic treatment effects. Without exogeneity, the assumption that all observed common factors of the regions account for all biases when comparing the targeted and control regions, the general method of treatment effect does not exist. Therefore, the two methods have focused on the macroeconomic scene where the individual effect of a targeted region is desired. These special cases include a targeted region where the treatment occurs somewhere in the timeline and a group of control regions that are related to the target region and are not affected by the treatment. In the attempt to lessen the time-average treatment effect, both the SCM and the PDCM adopt the difference-in-difference method and create an artificial control of the targeted region allowing for time-variant estimation of the treatment. The benefit of allowing for a time-variant estimation lies in the event that the effects of the treatment would change over time, shrinking, spreading, auto-regressive and etc. Common examples for these settings are the signing of treaties, passing laws, riots or civil wars, and tax or wage changes. The SCM and the PDCM offers different perspective in achieving such goals.

3-1. Synthetic Control Method

The synthetic control method (SCM) that was proposed initially by Abadie and Gardeazabel (2003) discussed the effect of riot on the GDP growth of Spanish Basque County. Based on the original idea, Abadie et al. (2010) improved further the method to discuss the effects of California's tobacco tax program, Proposition 99, on the yearly per capita tobacco sales. The authors concluded that per capita tobacco sales in California dropped since 1988 when the tobacco tax program was implemented. Moreover, the magnitude of the drop is greater than what was proposed in a previous study by Fitchtenberg and Glantz (2000). The SCM allows for determining an individual and time-variant treatment effect. The method is used on a specific case of data in which a region (the targeted region) is affected by the treatment for all time periods after the treatment-period, and all other regions (the control regions) are not affected by the treatment in every time period.

The SCM formulates a weighted sum of the control regions in order to simulate the behaviors of the targeted region. The SCM generates a weight (W) of the control regions using pre-treatment-period data in which the weighted sum of the control regions mimics the pre-treatment-period behavior of the targeted region. The weight (W) in the SCM is derived by minimizing

$$\|H_1 - H_0 W\| V \tag{3-1}$$

where V is some (r+M x r+M) symmetric and positive semidefinite matrix, $\mathrm{W} = \left(w_2, \ldots, w_{J+1}\right)^T$ with $w_j \geq 0\ \forall j = 2, \ldots, J+1$, $H_1 = \left(a_1{}^T, \overline{y}_1^{K_1}, \ldots, \overline{y}_1^{K_M}\right)^T$ is a vector that describes the pre-treatment-period characteristics of the targeted region, and

$$H_0 = \begin{bmatrix} \left(a_2{}^T, \overline{y}_2^{K_1}, \ldots, \overline{y}_2^{K_M}\right)^T \\ \vdots \\ \left(a_{J+1}{}^T, \overline{y}_{J+1}^{K_1}, \ldots, \overline{y}_{J+1}^{K_M}\right)^T \end{bmatrix}$$ where $K_l = \left(k_{l1}, \ldots, k_{lT_0}\right)^T, \mathrm{l} = 1, \ldots, \mathrm{M}, \sum_{i=1}^{T_0} k_{li} = 1$

are (T_0x 1) vectors that define the linear combination of pre-treatment-period data of the control regions. That is, $\overline{y}_i^{K_l} = \sum_{s=1}^{T_0} k_{ls} y_{is}$. The minimization of W in Equation (2-1) can be regarded as the minimization of the difference between the targeted and the control region behaviors of the pre-treatment-period data. In the empirical application of the SCM, W is chosen to minimize $\|H_1 - H_0 W\|\mathrm{V}$, where V is a positive definite and diagonal matrix. For every possible V there is a unique W^* that minimizes $\|H_1 - H_0 W\|\mathrm{V}$. Therefore, W is treated as W(V) and the minimization equation becomes $\|H_1 - H_0 W(V)\|\mathrm{V}$, which solely depends on V. Since the minimization process depends only on pre-treatment-period data, the V that minimizes the pre-treatment-period mean square deviation (MSD) is chosen as $V^*$and the minimizing weight $W^* = W(V^*)$ is then determined. Therefore, the artificial control (or the counterfactual) of the targeted region is derived with the sum of the post-treatment-period control regions weighted by W^*, and the treatment effect is obtained as the difference between the actual and counterfactual targeted region.

In this article, the simplified version of the SCM was used for a generalized application. The Matlab codes that Abadie et al. (2010) provided on their webpage were modified for implementing the SCM on generated data sets while the minimization process was kept the same. As in the empirical analysis conducted by the authors, the choice of V was restricted to a positive definite and diagonal matrix. For simplicity, K_l, the linear combination of y_{it}, was picked as y_{jt}; all pre-treatment-period data were included. The optimization process of picking V^* involved the use of Matlab function *fmincon* provided in the authors' codes. The initial values of V that were chosen as the diagonal terms were the normalized standard deviations of H_0, the control region behaviors of the pre-treatment-period data, and with other terms zero. The initial V was input into *fmincon* with the loss function, also provided in the authors' Matlab codes, to find the V^* that minimized pre-treatment-period MSD.

The SCM not only takes into account the observed common factors but also employs control region data. The observed common factors are chosen by researchers, while all control regions are included. Slight modifications were made to the more generalized method for simplified computation and choices of less subjectivity. The combination weight W of the control regions was then chosen to minimize the difference in the observed common factors and pre-treatment-period data between the targeted and control regions. The treatment effect is the post-treatment-period difference between the

targeted region and the counterfactual, which is created by the weighted sum of the control regions in the post-treatment-periods.

3-2. Panel-Data Counterfactual Method

The panel-data counterfactual method (PDCM) that was proposed by Hsiao et al. (2012) employs a simple method of finding the artificial control that does not require specified common factors. The authors demonstrated by using PDCM that there is no treatment effect on the GDP growth of Hong Kong in its political integration with China in 1997 but a positive treatment effect on the GDP growth in its economic integration in 2003. The PDCM, same as the SCM, was developed to estimate an individual and time-variant treatment effect. As with the SCM, the PDCM was designed for specific cases of data where a region (the targeted region) is affected by the treatment for all time periods after the treatment-period, and all other regions (the control regions) are not affect by the treatment in every time period.

The PDCM method is constructed to minimize

$$\mathrm{E}\left[\frac{1}{T_0}\left(\widetilde{y_1^0}-\tilde{e}\bar{\alpha}-Y\tilde{\alpha}^*\right)^T A\left(\widetilde{y_1^0}-\tilde{e}\bar{\alpha}-Y\tilde{\alpha}^*\right)\right] \tag{3-2}$$

For objective and simplified use, a special case of the general method proposed by the authors is employed. With the restriction of A = I, the simplified method allows for applying conveniently an ordinary least-square regression of the targeted region to the

control regions to generate the counterfactual combination in the pre-treatment periods. That is, $\widehat{y_{1t}^0} = \hat{\tilde{\alpha}} + \widehat{\tilde{\alpha}^{*T}}\tilde{y}_t$ where $\tilde{y}_t$ is the vector of control regions and $\hat{\tilde{\alpha}}$ and $\widehat{\tilde{\alpha}^{*T}}$ are the intersection and beta coefficients, respectively, of $\tilde{y}_t$ regressing on y_{it}^0, the targeted region, in the pre-treatment periods t=1…T_0. Therefore the differences $\widehat{\Delta_{it}} = \mathrm{y}_{1t} - \hat{\tilde{\alpha}} + \widehat{\tilde{\alpha}^{*T}}\tilde{y}_t$ are obtained for $\mathrm{t} = T_0 + 1, \dots, T$. The pre-treatment-period mean square deviation (MSD) therefore equals to $\sum_{T_0+1}^{T} \widehat{\Delta_{it}}^2$.

In the method, the specific control regions to be used are selected from a larger pool of related regions. This selection process involves selecting the regions that would create a counterfactual combination that best-fit the targeted region in the pre-treatment periods. The selection is done in two steps. In the first step, the regions that minimize pre-treatment-period MSDs while control for the number of control regions used are picked. The suitable numbers of control regions are selected in the second step using two different criteria, the Akaike Information Criterion (AIC),

$$\mathrm{AIC(p)} = T_0 \ln\left(\frac{e_0' e_0}{T_0}\right) + 2(p+2)$$

and the corrected Akaike Information Criterion (AICC).

$$\mathrm{AICC(p)} = \mathrm{AIC(p)} + \frac{2(p+2)(p+3)}{T_0 - (p+1) - 2}$$

In both criteria, p is the number of related regions included and e_0 is the least-square residual using the regions picked in the first step. AIC and AICC, in essence, give the mean squared deviation with a penalty term for the number regions used. Therefore, the

number with the lowest AIC or AICC is the number suitable to be used in the PDCM. Since AICC includes an additional penalty term, the number of regions used in AICC is lower than, if not equals to, that used in AIC. The maximum number of related regions (M) that is considered in this process has a significant impact in empirical use. Since the selection process is essentially a grid search process, the number of pre-treatment-period MSDs to be calculated is 2^M, which grows exponentially as M increases. Therefore, M is restricted in practical use to the computing resources limits.

The PDCM employs data of the control regions without using the common factors. This eliminates the need for picking the observed common factor and in turn increases the objectivity of the studies. Due to limitations in computing resources, the maximum number of regions under consideration must be restricted, which means that an existing huge pool of related regions must be narrowed down in empirical studies. In determining the specific control regions to be used from the pool of related regions, criteria AIC and AICC are used. As the control regions to be used could be different, the estimation of using AIC and AICC are denoted as PDCM-AIC and PDCM-AICC, respectively. The PDCMs obtain a collection of beta-coefficients from regressing the control regions on the targeted region in pre-treatment-periods. The beta-coefficients are used to construct a counterfactual of the targeted region using data from

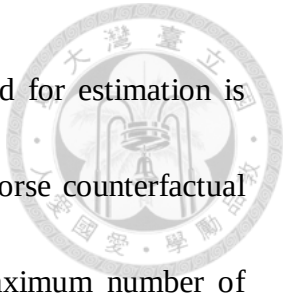
post-treatment-periods. The treatment effect is derived, as in the SCM, from the difference between the targeted region and the counterfactual.

4. Data Generation and Analysis

4-1. Monte Carlo Method and Data Generation

The objectives of both SCM and PDCM are to create a post-treatment-period counterfactual that mimics the behavior of the targeted region in the pre-treatment period. In the event that the treatment did not occur, the counterfactual generated by the two methods should be fitted as closely to the post-treatment period as possible. This closeness of fit is measured by the mean square deviation (MSD) between the counterfactual and the post-treatment-period targeted region. The smaller the MSD is, the more fit the counterfactual is. Since the counterfactuals of the SCM and the PDCM have been constructed differently, the need to develop a fitness test for comparing the two methods was warrant.

The treatment effect is not obtainable directly from the observational data, therefore the Monte Carlo method developed by Metropolis and Ulam (1949) suggested repeated simulation en route constructed data. Through simulations, the post-treatment MSD would reveal the closeness of fit of the methods. Since the desired MSD utilized of the unaffected data of the targeted and control regions, the constructed data need not specify the time, magnitude and type of change the treatment induced. By creating a data set that is structured in the same way throughout the timeline, the imaginary treatment period can be located individually on the same dataset. This way, when the treatment



period is closer to the starting period, the data that can be obtained for estimation is shorter and provide less information, leading an expectation of a worse counterfactual fit. Other expected factors in the counterfactual fit include the maximum number of related regions to be chosen as the control regions, the number of common factors included in the data generating model, and the number of observed factors among all common factors. If the control regions can be chosen from more related regions, the chance of a close counterfactual fit is higher. The greater the number of common factors is, the more complicated the data will be and the counterfactual is expected to be less likely to be a close fit. The difference in counterfactual construction enables information on observed factors to be taken into account in the SCM, and the fitness of the SCM counterfactual is expected to be better as more observed factors are added.

The difference of the SCM and the PDCM lies not only in the construction of the counterfactuals but also in the models the methods are constructed from. The model proposed by Hsiao et al. (2012) is designated in this study as Model (1), in which the common factors are fixed between regions and variant in time, and the factor-loadings are fixed in time and variant between regions.

$$y_{it}^0 = r_i + \lambda_i^T F_t + \varepsilon_{it}, \quad \text{i} = 1, \dots, \text{N}, \text{t} = 1, \dots, \text{T} \qquad \textbf{Model (1)}$$

where r_i is the individual fixed effect, λ_i^T is the (1 x K) unobserved factor loading vector where K is the number of factors, F_t is the (K x 1) vector of unobserved factors,

and ε_{it} is the idiosyncratic term with $\mathrm{E}(\varepsilon_{it}) = 0$. The model was generated by drawing r_i from N(0,1), λ_i^T from N(0,1), and ε_{it} from $\tau *$N(0,1). The factors were generated using AR(1),

$$F_t = \gamma * F_{t-1} + \epsilon_{it}$$

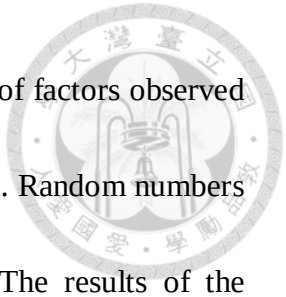
where γ is a (K x K) matrix with diagonal elements drawn from U(0,1) and every other elements 0. ϵ_{it} is drawn from N(0,1). τ is the mean of the diagonal elements of γ.

The data generating model Abadie et al. (2010) proposed was designated as Model (2), which includes common factors that are fixed in time and variant between regions, and factor-loadings that are fixed between regions and variant in time.

$$y_{it}^0 = r_t + \lambda^O{}_t F^O{}_i + \lambda^U{}_t F^U{}_i + \varepsilon_{it}\,, \quad \mathrm{i} = 1, \ldots, \mathrm{N}, \mathrm{t} = 1, \ldots, \mathrm{T} \qquad \textbf{Model (2)}$$

where r_t is the cross-unit-constant unknown common factor, $\lambda^O{}_t$ is the (1 x r) vector of unknown parameters, a_i is a (r x 1) observed common factors, $\lambda^U{}_t$ is a (1 x f) vector of unknown parameters, $F^U{}_i$ is a (f x 1) vector of unobserved common factors and ε_{it} is the idiosyncratic term with $\mathrm{E}(\varepsilon_{it}) = 0$. r + f = K is the total number of factors. r_t is generated from N(0,1), $\lambda^O{}_t$ and $\lambda^U{}_t$ from N(1,1), and ε_{it} from N(0,1). The common factors $F^O{}_i$ and $F^U{}_i$ are generated from $\chi^2(1)$.

Using Model (1) and Model (2), data were generated for simulations of the SCM and the two PDCM (PDCM-AIC and PDCM-AICC) methods. For each trial session that



employs different treatment period, number of factors used, number of factors observed and maximum number of related regions, 1000 trials were performed. Random numbers were independently drawn for every trial in every trial session. The results of the simulation as well as the random numbers were recorded for further investigation.

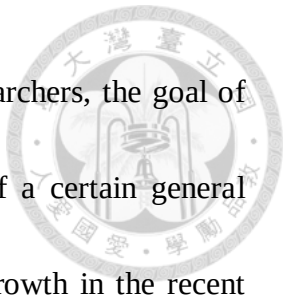
The data structure remained unchanged in all time periods of a single trial. As both artificial control methods aimed to replicate pre-treatment targeted region behavior under different environments, the uniformity of data structure provided the actual data of targeted region in absence of the treatment. Therefore, the difference between the uniform structure data and the estimates of the artificial methods would reveal their accuracy in estimations.

4-2. Data Processing Programs and Cross Checking

The SCM and two PDCMs with varying control region selection criteria were applied to the simulation data generated by Model (1) and Model (2). The results and the realization of the random numbers were recorded for analyses.

The MSD between the counterfactuals generated and the actual data of the targeted region were stored for direct comparison including: between SCM counterfactual and actual targeted region, between PDCM-AIC counterfactual and actual targeted region, and between PDCM-AICC counterfactual and actual targeted region.

The realization of random numbers in every trial was stored and processed for regression analysis of factors affecting the fitness of the counterfactuals. Specifically, the characteristics of random variables in the Model (1) includes the mean of the common factors' variance over time (mean_var_factor), the mean of the common factors' mean over time (mean_mean_factor), constant of targeted region (constant_y1), the mean and variance of the constants of control region (mean_constant_yn and var_constant_yn), the mean and variance of the factor-loadings of the targeted region (mean_fl_y1 and var_fl_y1), mean and variance of the factor-loadings of the control region's mean and variance across regions (mean_mean_fl_yn, var_mean_fl_yn, mean_var_fl_yn and var_var_fl_yn), mean of the epsilon of targeted region (mean_ep_y1), and the mean and variance of the mean of the control region's epsilon (mean_mean_ep_yn and var_mean_ep_yn). The characteristics of random numbers in the Model (2) were processed similarly, with changes to character factor and factor-loading due to the difference in data generating process. The characteristics of the common factors were split into those of the targeted region and those of the control regions. The characteristics of the factor-loadings were merged as one for the targeted region and the control regions. These characteristics were regressed on the 1000 trials in each session using Stata 12 (StataCorp, College Station, TX).



As the characteristics are in empirical sense unobservable to researchers, the goal of these analyses is to provide guidance only for the circumstance if a certain general impression is known. For example, the dataset of Taiwan's GDP growth in the recent decades may include a big variance in common factors as the price in Taipei grew rapidly and the prices in rural areas stagnated. Therefore, these basic characteristics would provide information on the goodness of fit of the counterfactual that is produced.

Besides the simulations on generated data, the performance of the applications of SCM and the two PDCM on real world data are of importance. The empirical data of California's cigarette consumption was estimated using the two PDCMs, and the empirical data of Hong Kong's political integration was estimated by the SCM. As the data of unaffected targeted region is unobservable, the goodness of fit cannot be determined if the results from distinct methods differ.

In 1988, the tax increase of 25 cents per package of cigarette, Proposition 99, was passed by voters in California. The SCM estimated the tobacco consumption reduction effect of the tax increase to be significantly larger than the previous study by Fichtenberg and Glantz (2000) suggests. Twenty percent of the accrued tax was allocated to an anti-tobacco educational program that included media advertisements, health and education budget, and promotion of clean air community. The Proposition 99 was widely believed to have decreased smoking in the state of California. The two

PDCMs were employed in cross checking this effect. The tobacco consumption data and common factors used were provided by Abadie et al. (2010) on their website (http://www.mit.edu/~jhainm/synthpage.html).

The other cross checking estimated the effect of Hong Kong's political integration on GDP growth. Hong Kong experienced a rapid growth under British sovereignty and was "returned" to China under "one country, two systems" on July 1997. The real GDP growth in Hong Kong dipped in 1997 but the Asian financial crisis in October and H5N1 Avian flu in December both broke out after the political integration. Therefore, the effect of the integration is difficult to be estimated. The PDCMs estimation suggested that the integration has no bearing on the real GDP growth of Hong Kong. The SCM was implemented to verify this result. All data used are obtained through OECD statistics, international financial statistics and CEIC database.

5. Results and Discussions

The result of implementing the PDCM on the data of California's Proposition 99 on cigarette tax is shown in Figure 1. The solid line represents the actual data of tobacco consumption in California over 1970 to 2000. Both the PDCM-AIC and the PDCM-AICC showed trends different from the results of the SCM. The PDCM-AIC estimation suggested that the Proposition 99 increases the tobacco sales; whereas the PDCM-AICC estimation showed that the Proposition 99 lowers the tobacco sales. The estimated effect is non-existent before year 1995 and far smaller than that of the SCM estimation. The result of implementing SCM on the data of Hong Kong's political integration with China is shown in Figure 2. The solid line represents the actual data of GDP growth in Hong Kong. The dashed line representing the SCM estimation showed the estimation difference between 1997 and 2002 and suggested that Hong Kong suffers a drop in GDP growth as a result of the political integration. This result is different from the estimations from the two PDCM methods which suggested the political integration has no bearing on the GDP growth. This cross checking result indicated that there are differences in the estimation of the two methods. Both Figures 1 and 2 showed that the estimations using the two different methods can indeed produce outcomes significantly different. From which, one would be led to totally different conclusions.

When a completely random white-noise data was implemented on both SCM and

PDCM methods (Figures 3 and 4). The fit is poor even for the pre-treatment-period data. However, when an accumulated white-noise data were implemented, the results are surprising. In some of the simulations, as shown in Figures 5 and 6, the estimation fits nicely before treatment-period and differs between the actual data and the estimations. If any conclusion could be drawn from these simulations, it points to a significant treatment effect, while there in fact is no change in the data structure. Therefore, the accumulated white-noise simulation study suggested that the two methods are not necessarily applicable with every data generating model.

The mean MSDs of 1000 trials of Model (1) structures with different settings are shown in Table 1. The majority of the mean MSDs from the SCM are larger than those from the PDCM-AICC, whereas the majority of the mean MSDs from the PDCM-AIC is slightly larger than those from the PDCM-AICC but still much smaller than those of the SCM. The settings where the mean MSDs of PDCMs were significantly larger were the ones that the number of total regions approached the number of pre-treatment-period. This could be explained by the estimation method of the PDCMs that involves regressing the control regions on the targeted region. When the number of total regions approached the number of pre-treatment-period, the regression was possibly over-fitted so that the post-treatment-period estimation fared poorly. Since the SCM method does not utilize such regression approach, their mean MSDs do not show this character. The

mean MSDs shown in Table 2 were from different settings using 1000 trials of Model (2). The results showed patterns identical to those on Table 1. Contrary to the hypothesis, the numbers of observed common factors do not affect the fit of the SCM. As the maximum number of regions can be subjectively changed by researchers, it therefore could be made lower when the number of pre-treatment-period is small. Therefore, the goodness of fit of the PDCMs is better than that of the SCM in empirical use.

The results of MSD decomposition are shown in Tables 3 to 14. Tables 3, 4 and 5 showed the results of decomposing the 1000 MSDs of Model (1) estimated by the SCM with 10, 20 and 30 pre-treatment-periods, respectively. The Mean_var_factor and Var_fl_y1 are the two variables with consistent significant and positive coefficients. Mean_var_factor represents the mean of the factors' variance over time and Var_fl_y1 is the variance of the factor-loadings of the targeted region. These results indicated that the magnitudes of both the factors and the factor-loadings of targeted region changed were inversely related to the accuracy of estimation. Tables 6 to 8 showed similar results for the PDCM-AICC estimations.

The variation of factors is a big issue for prediction because when the variations of factors are large overtime, the scale relationship between the factors may suffer a greater change, which in turn would result in a larger mean MSD. The reason that larger variations of targeted region's factor-loadings led to a reduced accuracy of estimation

may be that when a factor-loading is larger than the rest, the same change in the corresponding factor would result in more change to the region. This is a problem when the regions used are small because the prediction would be vulnerable to the sudden spike in that particular factor, and the small number of regions is not sufficient to cover the jump.

Tables 9 to 14 displayed the MSD decompositions of Model (2). The results were similar to those in the SCM and the PDCM-AICC estimation analyses as Var_factor_y1 was consistently significant and positive. Since Var_factor_y1 denotes the variance of factors of the targeted region, this result is consistent with the finding in Model (1).

Analysis of the simulations revealed that the PDCMs estimated the counterfactual more accurately than the SCM. The cross checking results showed that the two methods may estimate counterfactuals differently in empirical use, which may lead to different conclusions in the treatment effect studies. The methods, though showed expected unfitness to white noise data, were possibly misleading when faced with accumulated white noise data. With the exception of the cases where the maximum number of regions is closed to the number of treatment period, the PDCM-AICC clearly outperforms the SCM. The regressions analyses on MSDs showed that both methods estimated more accurately when the variances of common factors and factor-loadings were smaller.

6. Conclusion

The Monte Carlo results showed the estimation accuracy of the Panel-Data Counterfactual Method (PDCM) proposed by Hsiao et al. (2012) and the Synthetic Control Method (SCM) proposed by Abadie et al. (2012). The PDCM in most cases is more accurate in estimations than is the SCM, and the exceptions can be easily avoided. However, both the SCM and PDCM must be applied with caution because the results from the treatment could be misleading under certain settings such as data similar to accumulations of white noise. Both of the methods work better when the variation of factor in time and factor-loading across regions are smaller. Therefore, the methods may be more useful in topics that are known to be steadier and more balanced in factor influences, such as the manufacturing or dairy industries. The PDCM's robust performance under different model settings, together with easy applicability and improved credibility, make it a tool of choice in treatment effect studies.

Reference

Abadie, A., A. Diamond, and J. 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(490), 493-505.

Abadie, A., and J. Gardeazabal, (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93(1), 112-132.

Ashenfelter, O., (1978), "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, 60(1), 47-57.

Ashenfelter, O., and D. Card, (1985), "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs," *Review of Economics and Statistics*, 67(4), 648-660.

Card, D., (1990), "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43, 245-257.

Card, D., and A. Krueger, (1994), "Minimum Wages and Employment: A Case Study of the Fast-food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84 (4), 772-784.

Card, D., and D. Sullivan, (1988), "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment," *Econometrica*, 56(3), 497-530.

Fichtenberg, C. M., and S. A. Glantz, (2000), "Association of the California Tobacco Control Program with Declines in Cigarette Consumption and Mortality from Heart Disease," *New England Journal of Medicine*, 343(24), 1772-1777.

Fraker, T., and R. Maynard, (1987), "The Adequacy of Comparison Group Designs for Evaluations of Employment-Related Programs," *Journal of Human Resources*, 22(2), 194-227.

Heckman, J. J., and R. Robb, (1985), "Alternative Methods for Evaluating the Impact of Interventions: An Overview," *Journal of Econometrics*, 30(1), 239-267.

Holland, P., (1986), "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81(396), 945-960.

Hsiao, C., H. S. Ching, and S. K. Wan, (2012), "A Panel Data Approach for Program Evaluation: Measuring the Benefits of Political and Economic Integration of Hong Kong with Mainland China," *Journal of Applied Econometrics*, 27(5), 705-740.

Imbens, G. W., and J. M. Wooldridge, (2009), "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47(1), 5-86.

Lalonde, R. J., (1986), "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, 76(4), 604-620.

Manski, C. F., (1990), "Nonparametric Bounds on Treatment Effects," *American*

Economic Review, 80(2), 319-323.

Metropolis, N., and S. Ulam, (1949), "The Monte Carlo Method," *Journal of the American Statistical Association*, 44(247), 335-341.

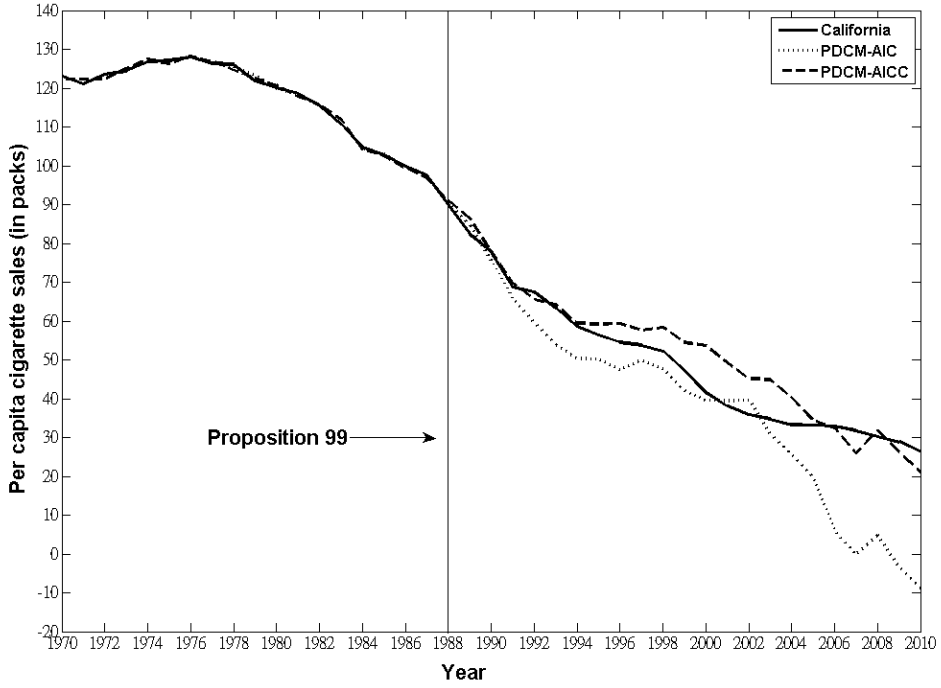


Figure 1. Trends in per-capita cigarette sales in California between 1970 and 2010 before and after the tax hike from the Proposition 99 in 1988. The solid line indicates the real data and the two dashed lines indicate counterfactuals by the panel-data counterfactual - Akaike Information Criterion method (PDCM-AIC) and – corrected Akaike Information Criterion method (PDCM-AICC), respectively.

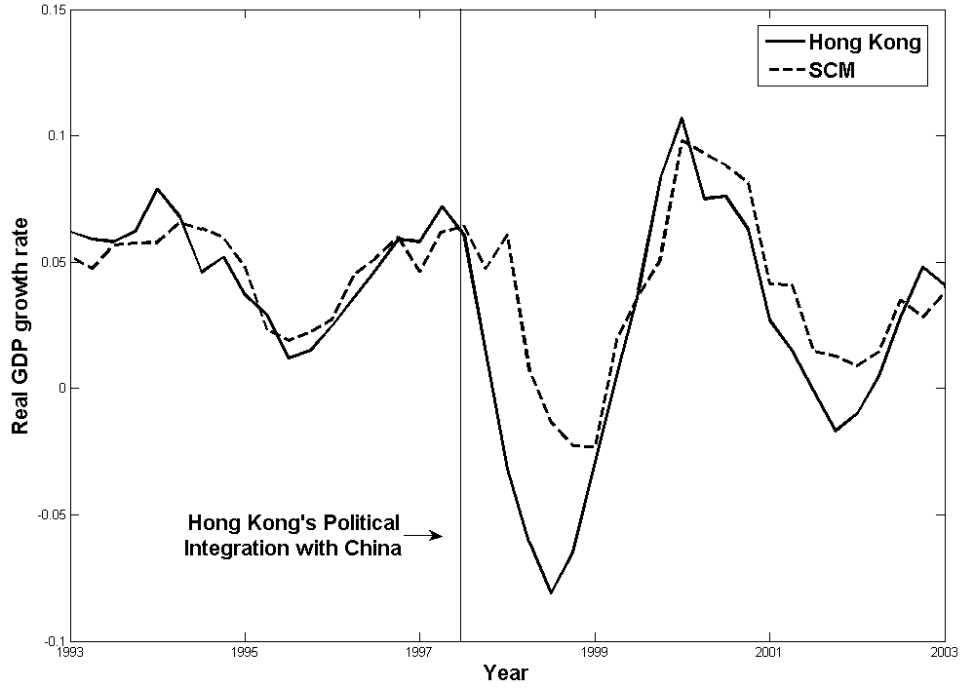


Figure 2. Trends in real GDP growth rate of Hong Kong between 1990 and 2003 before and after its political integration with China in 1997 and economic integration in 2003. The solid line denotes the real data and the dashed line the counterfactual by the he synthetic control method (SCM).

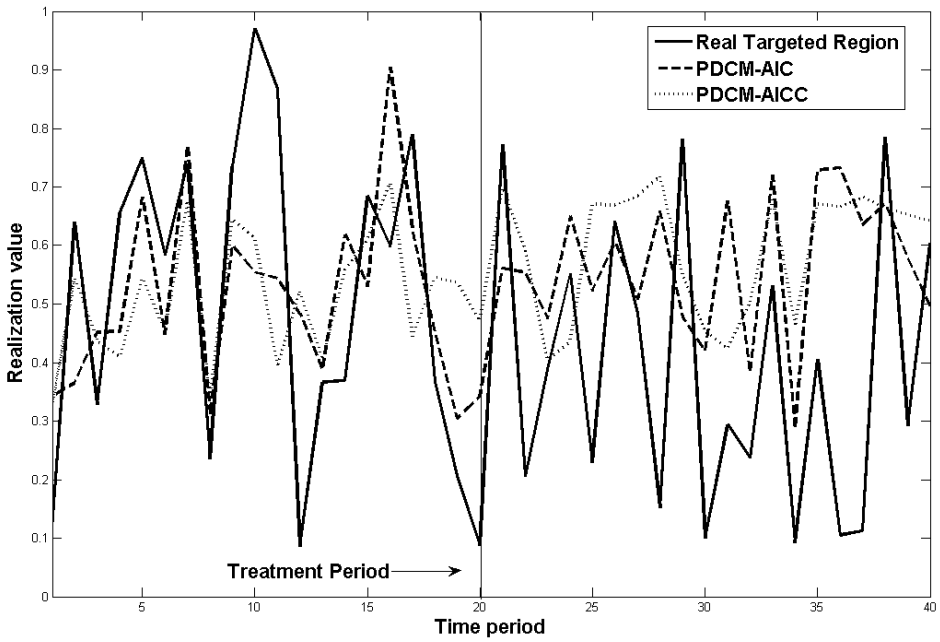


Figure 3. Counterfactual of white noise data estimated by the PDCMs: The solid line indicates the real targeted region, and the two dashed lines indicate counterfactuals by the panel-data counterfactual - Akaike Information Criterion method (PDCM-AIC) and – corrected Akaike Information Criterion method (PDCM-AICC), respectively.

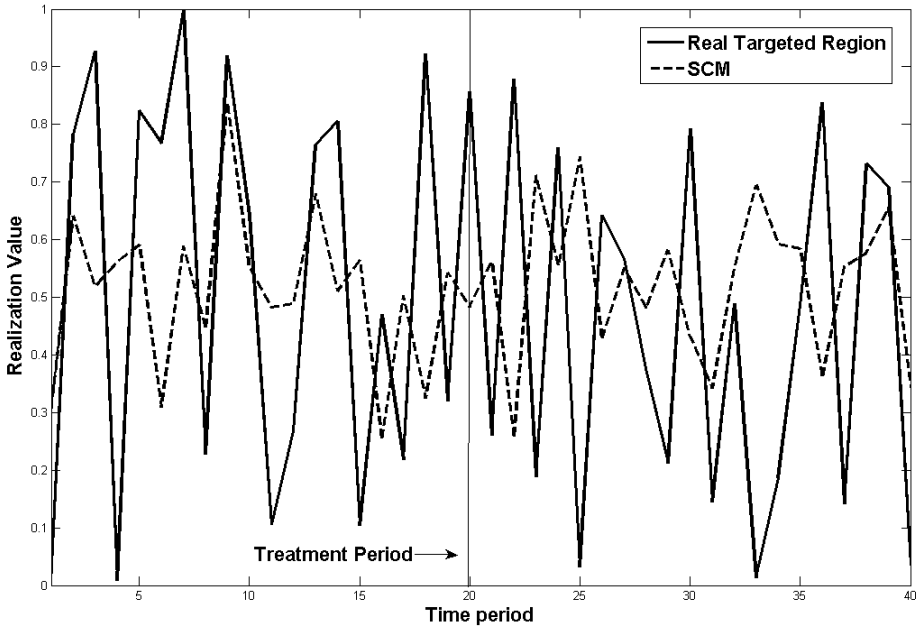


Figure 4. Counterfactual of white noise data estimated by the SCM: The solid line denotes the real targeted region and the dashed line the counterfactual by the synthetic control method (SCM).

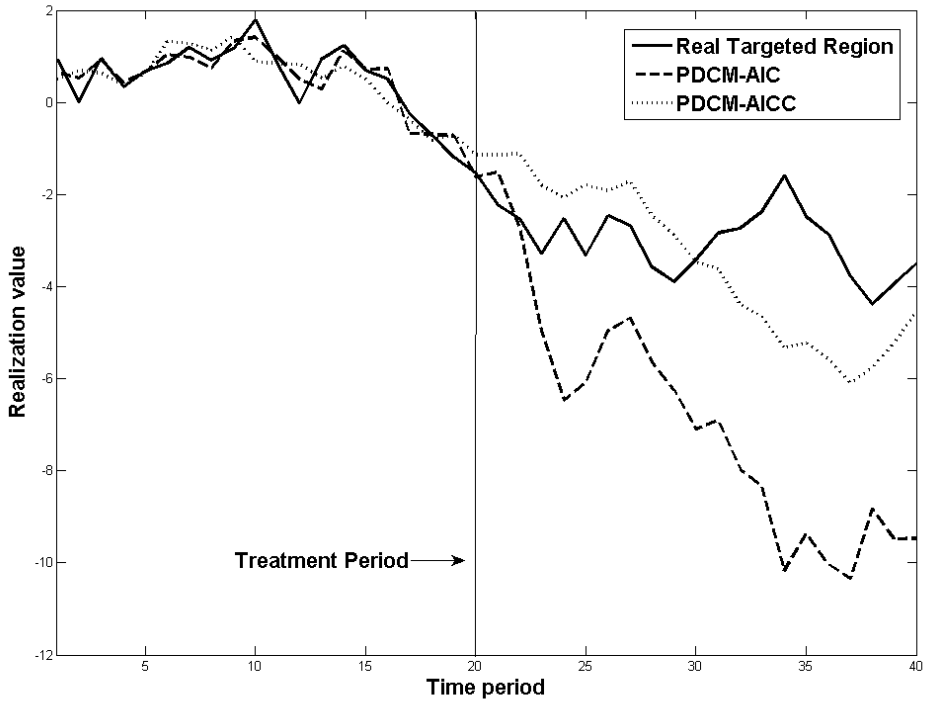


Figure 5. Counterfactual of the accumulations of white noise data estimated by the PDCMs: The solid line indicates the real targeted region, and the two dashed lines indicate counterfactuals by the panel-data counterfactual - Akaike Information Criterion method (PDCM-AIC) and – corrected Akaike Information Criterion method (PDCM-AICC), respectively.

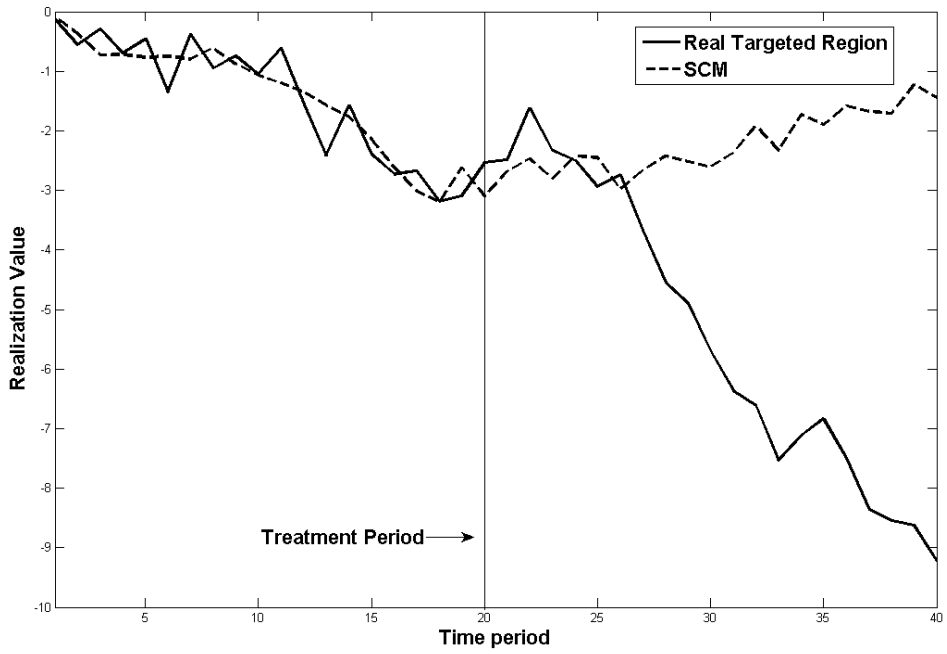


Figure 6. Counterfactual of the accumulations of white noise data estimated by the SCM: The solid line denotes the real targeted region and the dashed line the counterfactual by the synthetic control method (SCM).

Table 1. Mean Square Deviations (MSD)[1] of Counterfactual Estimates in Model (1) Generated Data

Variables[2]			Estimation Methods[3]		
Factor	Treatment Period	Region	PDCM-AIC	PDCM-AICC	SCM
2	10	4	2.057	2.305	5.706
2	10	6	2.034	1.931	3.406
2	10	8	3.278	1.544	3.080
2	10	10	32621.228	32621.228	2.418
2	20	4	1.270	1.353	5.435
2	20	6	0.869	0.878	3.605
2	20	8	0.917	0.909	3.000
2	20	10	0.760	0.699	2.082
4	10	4	6.841	7.534	12.014
4	10	6	6.651	5.817	8.307
4	10	8	7.518	4.546	7.047
4	10	10	44607.660	44607.660	6.340
4	20	4	4.438	4.581	10.680
4	20	6	2.305	2.620	7.834
4	20	8	1.540	1.689	6.382
4	20	10	1.322	1.315	5.183

[1] Mean of 1000 trials in each variable setting.
[2] Factor and region denote the total common factors used in data generation and the maximum possible numbers of regions included, respectively.
[3] The panel-data counterfactual method using Akaike information criterion (PDCM-AIC); the panel-data counterfactual method using corrected Akaike information criterion (PDCM-AICC); synthetic control method (SCM).

Table 2. Mean-Square-Deviations (MSD)[1] of Counterfactual Estimates in Model (2) Generated Data

Variables[2]			PDCM[3]		SCM[4]	
Factor	Treatment Period	Region	AIC	AICC	1	4
4	10	4	10.448	9.924	15.046	13.913
4	10	5	11.086	9.345	12.471	13.418
4	10	6	12.206	9.211	12.098	12.599
4	10	7	11.386	7.320	11.319	10.827
4	10	8	19.703	8.323	10.922	9.687
4	10	9	51.366	68.128	9.983	9.295
4	10	10	62602.027	62602.027	8.406	10.390
4	20	4	7.806	7.813	14.060	15.337
4	20	5	6.930	6.967	13.214	13.695
4	20	6	6.171	6.223	12.842	11.315
4	20	7	5.426	5.437	10.433	11.249
4	20	8	5.384	5.257	8.873	9.295
4	20	9	5.050	4.894	10.032	8.500
4	20	10	4.663	4.413	7.920	9.507
4	30	4	7.599	7.593	15.696	15.003
4	30	5	7.115	7.123	11.938	12.612
4	30	6	5.197	5.227	11.339	11.983
4	30	7	5.119	5.169	10.426	11.043
4	30	8	4.592	4.609	11.585	8.582
4	30	9	3.997	3.981	8.657	8.092
4	30	10	3.682	3.627	8.891	7.997

[1] Mean of 1000 trials in each variable setting.
[2] Factor and region denote the total common factors used in data generation and the maximum possible numbers of regions included, respectively.
[3] PDCM is the panel data counterfactual method. AIC denotes the method using Akaike information criterion. AICC denotes the method using corrected Akaike information criterion.
[4]SCM denotes the synthetic control method. Numbers 1 and 4 are the numbers of observed factors.

Table 3. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Synthetic Control Method in Model (1) Generated Data with Treatment Period at 10

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_var_factor	8.684*** (11.97)	7.384*** (11.84)	4.176*** (9.25)	4.204*** (9.22)	5.783*** (10.22)	4.746*** (9.95)	5.598*** (10.24)
Mean_mean_factor	0.706 (0.58)	-4.456*** (-4.31)	1.218 (1.70)	-1.230 (-1.51)	-2.748** (-3.19)	-0.354 (-0.47)	-3.596*** (-3.83)
Constant_y1	-0.0924 (-0.18)	-0.275 (-0.62)	-0.514 (-1.64)	0.214 (0.66)	0.170 (0.46)	0.279 (0.92)	1.094** (2.67)
Mean_constant_yn	0.543 (0.60)	-0.430 (-0.46)	0.436 (0.62)	0.119 (0.15)	1.030 (1.05)	0.700 (0.80)	-0.828 (-0.69)
Var_constant_yn	0.409 (0.52)	0.495 (0.64)	0.416 (0.74)	-0.322 (-0.52)	-0.0336 (-0.04)	0.650 (0.98)	0.145 (0.15)
Mean_fl_y1	1.817 (1.71)	0.781 (0.86)	-1.362* (-2.14)	0.561 (0.85)	0.917 (1.24)	0.221 (0.35)	-0.400 (-0.48)
Var_fl_y1	6.213*** (7.29)	10.08*** (13.73)	5.958*** (11.43)	5.773*** (11.56)	5.908*** (9.31)	4.847*** (9.80)	5.908*** (8.94)
Mean_mean_fl_yn	1.605 (0.86)	-1.039 (-0.58)	2.553 (1.82)	-1.057 (-0.67)	-1.580 (-0.83)	4.353** (2.59)	-2.058 (-0.82)
Var_mean_fl_yn	5.498* (2.15)	9.480** (3.18)	6.526* (2.48)	5.339 (1.70)	4.958 (1.15)	8.946* (2.12)	2.261 (0.40)
Mean_var_fl_yn	0.355 (0.17)	-0.918 (-0.46)	-1.171 (-0.85)	-3.151* (-2.23)	-0.817 (-0.45)	-1.506 (-0.95)	-3.133 (-1.49)
Var_var_fl_yn	-0.138 (-0.11)	-0.665 (-0.39)	1.723 (1.23)	0.304 (0.25)	-0.907 (-0.41)	5.477* (2.35)	-0.477 (-0.15)
Mean_ep_y1	3.432 (1.69)	-4.009* (-2.06)	0.337 (0.23)	-0.560 (-0.34)	-2.135 (-1.07)	2.218 (1.23)	1.115 (0.47)
Mean_mean_ep_yn	-1.624 (-0.47)	-2.395 (-0.65)	-0.102 (-0.03)	2.769 (0.69)	-8.550 (-1.62)	-6.394 (-1.25)	-5.238 (-0.70)
Var_mean_ep_yn	12.82 (1.20)	-6.196 (-0.51)	15.78 (1.71)	15.47 (1.26)	-16.19 (-0.98)	22.81 (1.37)	-7.154 (-0.28)
Tau1	1.476 (0.78)	-0.199 (-0.12)	2.448* (2.10)	0.0232 (0.02)	0.749 (0.52)	0.548 (0.46)	-4.654** (-3.01)
Tau2	0.749 (0.39)	2.848 (1.72)	0.755 (0.64)	-0.789 (-0.64)	2.376 (1.72)	1.358 (1.15)	0.923 (0.60)
Constant	-10.70*** (-4.89)	-11.81*** (-5.65)	-6.094*** (-4.16)	-1.960 (-1.20)	-7.437*** (-3.75)	-7.079*** (-3.95)	-2.797 (-1.20)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 4. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Synthetic Control Method in Model (1) Generated Data with Treatment Period at 20

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_var_factor	7.414***	6.484***	7.624***	4.626***	6.192***	2.803***	2.902***
	(10.40)	(11.07)	(8.33)	(9.58)	(10.11)	(9.50)	(10.18)
Mean_mean_factor	0.234	-0.725	-6.913***	1.034	-3.626***	2.015***	0.0839
	(0.19)	(-0.85)	(-4.56)	(1.27)	(-3.55)	(3.94)	(0.18)
Constant_y1	-0.208	-0.0609	-0.110	0.411	0.0801	-0.0245	-0.116
	(-0.43)	(-0.18)	(-0.18)	(1.25)	(0.18)	(-0.11)	(-0.64)
Mean_constant_yn	2.217**	0.633	2.536	0.346	-0.450	-0.511	-0.914
	(2.64)	(0.89)	(1.82)	(0.42)	(-0.38)	(-0.81)	(-1.55)
Var_constant_yn	0.227	-0.120	-0.221	-0.983	-0.0483	-0.612	0.117
	(0.30)	(-0.21)	(-0.20)	(-1.55)	(-0.06)	(-1.33)	(0.28)
Mean_fl_y1	0.387	-0.316	2.152	-0.945	0.513	-0.0610	-0.127
	(0.41)	(-0.46)	(1.77)	(-1.47)	(0.59)	(-0.14)	(-0.34)
Var_fl_y1	6.748***	5.702***	8.297***	5.362***	5.962***	5.167***	3.598***
	(8.69)	(10.01)	(8.29)	(9.64)	(8.62)	(14.97)	(11.50)
Mean_mean_fl_yn	-2.730	-2.190	-4.003	3.314*	-2.017	0.295	1.483
	(-1.57)	(-1.56)	(-1.54)	(2.03)	(-0.83)	(0.25)	(1.30)
Var_mean_fl_yn	4.485	10.01***	0.220	7.893*	3.660	8.022**	4.283
	(1.93)	(4.39)	(0.04)	(2.42)	(0.73)	(2.76)	(1.60)
Mean_var_fl_yn	-7.224***	-1.276	-4.331	-2.102	-4.687*	-0.588	-0.869
	(-3.58)	(-0.86)	(-1.70)	(-1.38)	(-2.34)	(-0.55)	(-0.89)
Var_var_fl_yn	2.101	0.932	1.489	-0.316	2.596	-0.130	0.523
	(1.45)	(0.75)	(0.76)	(-0.19)	(1.39)	(-0.09)	(0.35)
Mean_ep_y1	-2.425	0.941	3.015	-1.498	0.427	-0.894	2.559*
	(-1.25)	(0.64)	(1.06)	(-0.87)	(0.18)	(-0.67)	(2.23)
Mean_mean_ep_yn	-7.315*	-1.867	7.014	-11.14**	12.47	-0.382	-0.306
	(-2.22)	(-0.64)	(1.10)	(-2.69)	(1.95)	(-0.11)	(-0.09)
Var_mean_ep_yn	10.03	-16.03	109.9***	20.95	-24.46	17.04	-5.681
	(1.12)	(-1.75)	(5.95)	(1.61)	(-1.24)	(1.51)	(-0.51)
Tau1	0.767	1.233	-0.421	-1.079	0.881	0.675	1.469*
	(0.42)	(0.99)	(-0.19)	(-0.89)	(0.55)	(0.82)	(2.05)
Tau2	4.374*	3.878**	-5.901**	-0.0471	3.175	0.779	1.326
	(2.39)	(3.10)	(-2.64)	(-0.04)	(1.94)	(0.96)	(1.87)
Constant	-6.852***	-8.876***	-8.673**	-3.272	-6.576**	-3.784**	-3.331**
	(-3.35)	(-5.56)	(-2.98)	(-1.89)	(-2.91)	(-3.21)	(-3.11)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 5. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Synthetic Control Method in Model (1) Generated Data with Treatment Period at 30

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_var_factor	5.877***	10.60***	8.600***	7.362***	2.816***	1.386**	2.613***
	(8.15)	(9.70)	(11.21)	(9.88)	(6.75)	(2.72)	(6.23)
Mean_mean_factor	0.756	-0.868	0.664	4.045***	-0.0997	-0.241	-1.462*
	(0.70)	(-0.51)	(0.48)	(3.35)	(-0.13)	(-0.26)	(-2.24)
Constant_y1	0.591	-0.0820	-0.254	0.573	0.0476	0.383	-0.286
	(1.29)	(-0.12)	(-0.44)	(1.23)	(0.17)	(1.01)	(-1.08)
Mean_constant_yn	0.402	-0.470	-0.629	0.425	-0.665	-0.676	-0.514
	(0.50)	(-0.33)	(-0.50)	(0.37)	(-0.88)	(-0.62)	(-0.65)
Var_constant_yn	-0.272	0.661	-0.256	-0.931	-0.0457	-0.748	-0.967
	(-0.37)	(0.59)	(-0.23)	(-1.07)	(-0.08)	(-0.92)	(-1.66)
Mean_fl_y1	-0.847	0.678	-1.201	-0.853	0.324	0.279	-0.485
	(-0.94)	(0.49)	(-1.07)	(-0.94)	(0.53)	(0.36)	(-0.92)
Var_fl_y1	5.842***	8.245***	5.309***	4.951***	5.943***	6.386***	4.629***
	(8.00)	(7.53)	(5.82)	(6.31)	(12.29)	(10.40)	(10.64)
Mean_mean_fl_yn	-0.751	-1.880	-2.209	3.176	-0.577	-0.609	0.172
	(-0.48)	(-0.69)	(-0.90)	(1.39)	(-0.37)	(-0.29)	(0.11)
Var_mean_fl_yn	5.282*	3.922	6.122	9.791*	0.729	3.096	-0.201
	(2.53)	(0.84)	(1.32)	(2.08)	(0.23)	(0.66)	(-0.05)
Mean_var_fl_yn	0.308	-8.573**	-1.850	-2.744	-3.794**	-3.549	-2.083
	(0.16)	(-3.03)	(-0.73)	(-1.25)	(-2.71)	(-1.82)	(-1.53)
Var_var_fl_yn	-0.593	3.479	1.199	2.101	1.386	-0.371	0.349
	(-0.45)	(1.50)	(0.45)	(0.87)	(0.70)	(-0.13)	(0.21)
Mean_ep_y1	-2.092	0.817	-3.363	3.586	-5.791***	-6.266**	-0.640
	(-1.18)	(0.28)	(-1.24)	(1.53)	(-3.57)	(-2.70)	(-0.40)
Mean_mean_ep_yn	4.170	-1.648	4.333	2.355	-1.274	-1.414	1.470
	(1.37)	(-0.29)	(0.73)	(0.40)	(-0.30)	(-0.22)	(0.31)
Var_mean_ep_yn	29.76***	-11.53	62.77***	-4.095	13.15	131.8***	5.748
	(3.40)	(-0.67)	(3.50)	(-0.23)	(0.95)	(5.86)	(0.34)
Tau1	0.436	-5.859*	-2.987	1.238	1.531	0.477	1.803
	(0.25)	(-2.28)	(-1.47)	(0.71)	(1.41)	(0.33)	(1.84)
Tau2	0.126	-0.635	2.946	-3.307	0.834	-1.117	-0.261
	(0.07)	(-0.25)	(1.40)	(-1.88)	(0.77)	(-0.77)	(-0.26)
Constant	-6.928***	-6.644*	-12.32***	-6.183*	-2.404	-1.418	-1.000
	(-3.50)	(-2.06)	(-4.47)	(-2.55)	(-1.58)	(-0.68)	(-0.64)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 6. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Panel-Data Counterfactual Method using Corrected Akaike Information Criterion in Model (1) Generated Data with Treatment Period at 10

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_var_factor	5.742*** (12.45)	5.325*** (10.50)	3.460*** (6.30)	2.461*** (7.00)	3.530*** (8.52)	2.317 (0.25)	-14621.7 (-0.27)
Mean_mean_factor	-1.393 (-1.79)	-0.816 (-0.97)	0.138 (0.16)	0.973 (1.55)	-4.004*** (-6.35)	21.95 (1.54)	41883.5 (0.45)
Constant_y1	0.688* (2.10)	-0.420 (-1.17)	0.878* (2.30)	-0.110 (-0.44)	0.191 (0.71)	2.275 (0.39)	-18869.9 (-0.47)
Mean_constant_yn	-0.634 (-1.10)	1.197 (1.57)	-0.212 (-0.25)	-0.233 (-0.39)	-0.515 (-0.72)	12.78 (0.76)	-12156.1 (-0.10)
Var_constant_yn	-0.227 (-0.45)	-0.329 (-0.53)	1.065 (1.57)	0.589 (1.23)	0.593 (1.06)	16.09 (1.27)	-78818.8 (-0.86)
Mean_fl_y1	-0.869 (-1.28)	0.264 (0.36)	-0.426 (-0.55)	0.131 (0.26)	-0.0831 (-0.15)	-13.44 (-1.12)	68469.1 (0.83)
Var_fl_y1	6.473*** (11.94)	5.132*** (8.60)	5.068*** (7.99)	2.500*** (6.50)	3.853*** (8.29)	5.531 (0.59)	-14651.7 (-0.22)
Mean_mean_fl_yn	-0.226 (-0.19)	1.149 (0.79)	-2.881 (-1.69)	1.312 (1.08)	1.291 (0.93)	-24.31 (-0.76)	-611565.9* (-2.47)
Var_mean_fl_yn	-0.261 (-0.16)	-0.710 (-0.29)	-3.289 (-1.03)	-4.242 (-1.75)	-3.247 (-1.03)	-47.21 (-0.59)	-557375.0 (-1.01)
Mean_var_fl_yn	-1.897 (-1.41)	-2.146 (-1.33)	-4.858** (-2.91)	-0.791 (-0.73)	-1.404 (-1.06)	-34.93 (-1.16)	87950.6 (0.42)
Var_var_fl_yn	1.023 (1.29)	1.898 (1.39)	4.341* (2.55)	-0.399 (-0.42)	1.104 (0.67)	37.62 (0.85)	-246551.6 (-0.80)
Mean_ep_y1	1.549 (1.20)	1.997 (1.26)	-2.151 (-1.19)	-2.861* (-2.28)	-2.217 (-1.52)	18.11 (0.53)	-171458.8 (-0.73)
Mean_mean_ep_yn	-0.0374 (-0.02)	-1.086 (-0.36)	5.219 (1.24)	1.169 (0.38)	8.241* (2.13)	49.98 (0.51)	-1401605.4 (-1.89)
Var_mean_ep_yn	14.43* (2.13)	9.158 (0.93)	32.95** (2.94)	31.41*** (3.32)	9.998 (0.83)	268.7 (0.85)	1088248.8 (0.44)
Tau1	0.132 (0.11)	1.497 (1.10)	2.517 (1.78)	2.669** (2.83)	2.431* (2.31)	22.62 (0.99)	-71223.9 (-0.47)
Tau2	0.134 (0.11)	0.730 (0.54)	0.0863 (0.06)	1.616 (1.70)	1.318 (1.30)	16.39 (0.73)	35238.8 (0.23)
Constant	-6.567*** (-4.72)	-6.502*** (-3.82)	-3.724* (-2.09)	-3.654** (-2.91)	-5.432*** (-3.74)	3.625 (0.11)	148610.5 (0.65)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 7. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Panel-Data Counterfactual Method using Corrected Akaike Information Criterion in Model (1) Generated Data with Treatment Period at 20

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_var_factor	2.873*** (8.29)	1.664*** (6.16)	2.551*** (9.70)	1.053*** (8.65)	1.399*** (7.84)	0.429*** (5.78)	0.538*** (6.40)
Mean_mean_factor	-0.292 (-0.49)	-0.938* (-2.38)	-1.475*** (-3.38)	0.367 (1.79)	0.702* (2.36)	0.151 (1.17)	0.0638 (0.47)
Constant_y1	-0.163 (-0.69)	0.159 (1.00)	0.224 (1.26)	0.104 (1.25)	0.153 (1.20)	0.0900 (1.62)	0.0825 (1.53)
Mean_constant_yn	0.484 (1.19)	-0.257 (-0.78)	0.140 (0.35)	0.193 (0.92)	0.154 (0.45)	0.0395 (0.25)	0.314 (1.80)
Var_constant_yn	-0.221 (-0.60)	-0.00651 (-0.03)	-0.244 (-0.78)	-0.0131 (-0.08)	0.104 (0.41)	-0.153 (-1.33)	-0.0494 (-0.40)
Mean_fl_y1	0.573 (1.25)	-0.156 (-0.49)	-0.277 (-0.79)	0.200 (1.23)	-0.165 (-0.65)	-0.0707 (-0.64)	0.0452 (0.41)
Var_fl_y1	4.856*** (12.85)	3.117*** (11.86)	2.808*** (9.76)	1.013*** (7.22)	1.037*** (5.15)	0.716*** (8.25)	0.718*** (7.78)
Mean_mean_fl_yn	-1.069 (-1.26)	0.124 (0.19)	-0.441 (-0.59)	0.404 (0.98)	-0.339 (-0.48)	-0.0308 (-0.10)	-0.670* (-1.99)
Var_mean_fl_yn	-0.578 (-0.51)	-0.258 (-0.25)	-2.177 (-1.51)	-2.585** (-3.13)	1.295 (0.89)	-0.137 (-0.19)	-0.263 (-0.33)
Mean_var_fl_yn	-2.320* (-2.36)	-1.199 (-1.75)	-2.049** (-2.80)	-0.691 (-1.79)	-1.018 (-1.74)	-0.829** (-3.08)	-0.815** (-2.81)
Var_var_fl_yn	-0.00451 (-0.01)	1.640** (2.85)	0.672 (1.20)	0.0469 (0.11)	1.835*** (3.39)	0.439 (1.19)	0.528 (1.19)
Mean_ep_y1	1.036 (1.10)	0.127 (0.19)	1.353 (1.65)	0.324 (0.75)	0.753 (1.07)	-0.102 (-0.30)	0.533 (1.57)
Mean_mean_ep_yn	-2.076 (-1.29)	1.056 (0.78)	3.859* (2.11)	-0.958 (-0.92)	-0.336 (-0.18)	2.121* (2.35)	1.312 (1.33)
Var_mean_ep_yn	6.007 (1.38)	4.726 (1.12)	22.99*** (4.33)	6.891* (2.10)	16.08** (2.80)	16.48*** (5.80)	13.13*** (3.97)
Tau1	1.141 (1.30)	1.727** (2.99)	0.290 (0.45)	0.689* (2.25)	0.306 (0.66)	0.607** (2.95)	0.521* (2.46)
Tau2	1.031 (1.16)	1.048 (1.82)	0.000128 (0.00)	1.161*** (3.77)	0.749 (1.58)	0.770*** (3.77)	0.680** (3.25)
Constant	-3.343*** (-3.36)	-3.003*** (-4.07)	-2.717** (-3.24)	-0.959* (-2.19)	-2.139** (-3.25)	-0.167 (-0.57)	-0.302 (-0.95)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 8. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Panel-Data Counterfactual Method using Corrected Akaike Information Criterion in Model (1) Generated Data with Treatment Period at 30

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_var_factor	1.436*** (5.94)	1.086*** (5.59)	0.704*** (7.18)	0.548*** (3.46)	0.601*** (8.35)	0.266*** (4.17)	0.307*** (5.68)
Mean_mean_factor	0.284 (0.78)	0.872** (2.89)	-0.108 (-0.61)	0.0918 (0.36)	0.156 (1.22)	0.297* (2.56)	-0.148 (-1.77)
Constant_y1	-0.0276 (-0.18)	0.147 (1.25)	0.0415 (0.57)	-0.00305 (-0.03)	0.0926 (1.86)	0.0419 (0.88)	-0.0224 (-0.66)
Mean_constant_yn	-0.191 (-0.71)	-0.0714 (-0.28)	-0.0344 (-0.21)	0.0624 (0.25)	-0.0591 (-0.45)	0.0780 (0.57)	-0.0143 (-0.14)
Var_constant_yn	-0.0862 (-0.35)	0.0609 (0.31)	-0.0477 (-0.34)	0.288 (1.55)	-0.0555 (-0.53)	0.0973 (0.96)	0.0383 (0.51)
Mean_fl_y1	-0.484 (-1.61)	-0.0723 (-0.29)	-0.142 (-0.99)	0.135 (0.70)	-0.174 (-1.66)	-0.0165 (-0.17)	-0.0648 (-0.96)
Var_fl_y1	3.007*** (12.28)	1.759*** (9.04)	1.201*** (10.32)	1.263*** (7.58)	0.778*** (9.32)	0.604*** (7.83)	0.479*** (8.57)
Mean_mean_fl_yn	-0.806 (-1.55)	-0.0155 (-0.03)	-0.302 (-0.96)	-0.0954 (-0.20)	0.0742 (0.27)	-0.280 (-1.06)	0.327 (1.56)
Var_mean_fl_yn	-1.162 (-1.66)	-0.0480 (-0.06)	-0.740 (-1.25)	-1.931 (-1.94)	-0.601 (-1.12)	-0.650 (-1.10)	-0.120 (-0.22)
Mean_var_fl_yn	-1.167 (-1.81)	-0.848 (-1.68)	-1.292*** (-3.97)	0.351 (0.75)	-0.679** (-2.81)	-0.863*** (-3.53)	-0.837*** (-4.77)
Var_var_fl_yn	0.537 (1.21)	0.538 (1.31)	0.484 (1.42)	-0.451 (-0.88)	0.848* (2.48)	0.554 (1.53)	0.270 (1.24)
Mean_ep_y1	-0.206 (-0.35)	-0.581 (-1.13)	-0.550 (-1.58)	-0.785 (-1.57)	-0.217 (-0.78)	-0.373 (-1.28)	-0.149 (-0.73)
Mean_mean_ep_yn	2.643** (2.58)	-1.137 (-1.13)	-0.326 (-0.43)	-2.313 (-1.85)	-0.906 (-1.25)	-0.788 (-1.00)	0.201 (0.33)
Var_mean_ep_yn	5.080 (1.73)	6.759* (2.22)	4.976* (2.17)	5.072 (1.34)	5.694* (2.39)	12.17*** (4.31)	12.85*** (5.95)
Tau1	0.974 (1.69)	1.255** (2.75)	0.834** (3.22)	0.595 (1.60)	0.732*** (3.89)	0.784*** (4.34)	0.361** (2.86)
Tau2	0.687 (1.20)	0.587 (1.32)	1.403*** (5.21)	0.570 (1.52)	0.764*** (4.08)	0.751*** (4.11)	0.306* (2.34)
Constant	-1.250 (-1.88)	-1.399* (-2.44)	-0.546 (-1.55)	-1.411** (-2.74)	-0.838** (-3.18)	-0.359 (-1.37)	0.106 (0.53)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 9. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Synthetic Control Method in Model (2) Generated Data with Treatment Period at 10

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_factor_y1	0.719 (1.39)	0.578 (1.26)	-0.284 (-0.82)	1.403*** (4.00)	-1.063*** (-3.30)	-0.463 (-1.58)	0.522* (2.07)
Var_factor_y1	3.969*** (9.85)	4.247*** (11.61)	2.902*** (10.41)	3.100*** (11.34)	2.990*** (11.80)	2.747*** (12.20)	2.713*** (13.67)
Mean_var_factor_yn	1.825* (2.50)	1.524* (2.09)	0.443 (0.69)	0.227 (0.34)	0.0248 (0.04)	0.0662 (0.10)	-0.0879 (-0.15)
Mean_mean_factor_yn	0.503 (0.58)	-0.240 (-0.26)	-0.122 (-0.16)	0.532 (0.63)	0.709 (0.87)	1.393 (1.73)	-0.205 (-0.30)
Constant_y1	-0.134 (-0.54)	0.274 (1.22)	-0.0639 (-0.37)	0.0293 (0.17)	0.0800 (0.50)	0.00556 (0.04)	0.299* (2.56)
Mean_constant_yn	0.0660 (0.15)	0.172 (0.38)	-0.0809 (-0.21)	0.00947 (0.02)	-0.120 (-0.29)	-0.0611 (-0.15)	0.189 (0.54)
Var_constant_yn	-0.206 (-0.57)	-0.629 (-1.67)	-0.627* (-2.07)	-0.00872 (-0.03)	0.258 (0.82)	-0.281 (-0.87)	-0.282 (-1.05)
Mean_mean_fl	5.358 (1.72)	3.252 (1.11)	3.379 (1.51)	0.748 (0.34)	-0.0435 (-0.02)	-0.801 (-0.46)	0.514 (0.34)
Var_mean_fl	2.798 (0.17)	4.419 (0.29)	-2.565 (-0.21)	-18.51 (-1.60)	4.790 (0.47)	-9.057 (-1.01)	2.878 (0.38)
Mean_var_fl	9.364*** (3.85)	0.637 (0.31)	7.336*** (4.54)	1.689 (1.05)	3.531* (2.43)	2.435 (1.89)	0.276 (0.25)
Var_var_fl	-4.512 (-0.57)	0.586 (0.08)	-8.624 (-1.57)	4.250 (0.74)	-7.537 (-1.63)	-7.612 (-1.56)	-0.113 (-0.03)
Mean_ep_y1	-2.170 (-1.40)	1.940 (1.38)	-2.890** (-2.65)	2.372* (2.20)	-0.239 (-0.25)	-0.222 (-0.25)	-0.238 (-0.31)
Mean_mean_ep_yn	1.420 (0.52)	-3.852 (-1.31)	-4.917 (-1.89)	-1.881 (-0.70)	-1.130 (-0.44)	-0.773 (-0.30)	3.387 (1.50)
Var_mean_ep_yn	-13.01 (-0.88)	-5.205 (-0.37)	16.57 (1.34)	-9.275 (-0.73)	-3.870 (-0.32)	22.07 (1.69)	-14.96 (-1.42)
Constant	-8.337 (-1.64)	1.654 (0.34)	2.230 (0.56)	1.075 (0.28)	1.545 (0.42)	3.044 (0.84)	-0.877 (-0.28)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 10. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Synthetic Control Method in Model (2) Generated Data with Treatment Period at 20

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_factor_y1	-1.826*** (-3.33)	-0.753 (-1.69)	0.563 (1.28)	0.137 (0.47)	0.106 (0.36)	-0.488* (-1.99)	-0.357 (-1.42)
Var_factor_y1	2.552*** (5.48)	2.557*** (7.22)	3.515*** (10.03)	2.721*** (11.96)	2.317*** (9.94)	2.491*** (13.62)	2.376*** (11.08)
Mean_var_factor_yn	3.245*** (4.26)	1.107 (1.59)	1.268 (1.56)	0.450 (0.76)	0.518 (0.82)	-0.973 (-1.73)	-0.434 (-0.70)
Mean_mean_factor_yn	-0.0455 (-0.05)	-0.372 (-0.43)	-1.635 (-1.66)	-0.596 (-0.84)	-0.518 (-0.64)	0.529 (0.77)	0.564 (0.74)
Constant_y1	-0.637* (-2.37)	0.0395 (0.19)	0.0971 (0.45)	-0.147 (-1.00)	0.186 (1.24)	0.0715 (0.58)	0.124 (0.96)
Mean_constant_yn	0.0461 (0.10)	0.0816 (0.19)	-0.436 (-0.86)	-0.0916 (-0.26)	0.469 (1.15)	0.499 (1.48)	0.296 (0.76)
Var_constant_yn	-0.201 (-0.52)	0.155 (0.46)	-0.128 (-0.32)	-0.0524 (-0.20)	-0.488 (-1.52)	-0.409 (-1.65)	-0.194 (-0.71)
Mean_mean_fl	-1.978 (-0.57)	-2.065 (-0.80)	2.731 (0.98)	0.566 (0.31)	1.162 (0.60)	-2.022 (-1.37)	2.278 (1.34)
Var_mean_fl	24.89 (1.45)	-9.612 (-0.68)	24.36 (1.75)	-9.166 (-0.90)	9.034 (0.94)	6.893 (0.87)	0.188 (0.02)
Mean_var_fl	3.846 (1.51)	7.967*** (3.91)	3.171 (1.60)	1.633 (1.22)	2.819* (2.01)	1.350 (1.22)	0.541 (0.45)
Var_var_fl	6.200 (0.70)	1.391 (0.20)	-0.198 (-0.03)	3.879 (0.81)	-2.140 (-0.40)	7.487* (1.97)	0.588 (0.13)
Mean_ep_y1	1.807 (1.04)	-1.064 (-0.81)	-0.698 (-0.50)	-0.179 (-0.20)	0.352 (0.36)	-1.002 (-1.28)	0.153 (0.18)
Mean_mean_ep_yn	-0.0749 (-0.03)	-1.673 (-0.63)	-0.283 (-0.09)	1.858 (0.86)	-2.373 (-0.93)	2.097 (1.03)	0.0579 (0.02)
Var_mean_ep_yn	-25.37 (-1.52)	-20.71 (-1.52)	-15.15 (-0.96)	0.318 (0.03)	3.894 (0.32)	1.680 (0.17)	3.804 (0.33)
Constant	1.543 (0.27)	1.800 (0.39)	-1.506 (-0.31)	-0.337 (-0.10)	0.880 (0.24)	2.416 (0.82)	-0.125 (-0.04)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 11. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Synthetic Control Method in Model (2) Generated Data with Treatment Period at 30

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_factor_y1	-1.415*	-0.777	-0.873	-0.973**	0.130	1.390***	-0.265
	(-2.48)	(-1.64)	(-1.91)	(-2.60)	(0.42)	(4.10)	(-0.92)
Var_factor_y1	3.547***	2.604***	1.386***	3.300***	2.304***	3.145***	2.891***
	(7.47)	(6.40)	(3.85)	(11.52)	(8.73)	(10.75)	(12.07)
Mean_var_factor_yn	0.628	0.687	0.152	-1.470	-0.885	0.302	-0.895
	(0.79)	(0.86)	(0.18)	(-1.92)	(-1.30)	(0.40)	(-1.30)
Mean_mean_factor_yn	1.411	1.788	0.211	0.855	1.131	-0.625	-0.100
	(1.44)	(1.84)	(0.21)	(0.92)	(1.30)	(-0.67)	(-0.12)
Constant_y1	-0.316	-0.434	0.182	-0.181	0.298	0.160	-0.0268
	(-1.10)	(-1.78)	(0.77)	(-0.94)	(1.92)	(0.97)	(-0.19)
Mean_constant_yn	-0.0925	-0.161	-0.700	0.218	-0.305	0.118	-0.111
	(-0.18)	(-0.33)	(-1.33)	(0.50)	(-0.73)	(0.25)	(-0.27)
Var_constant_yn	-0.108	-0.321	0.192	-0.551	-0.0756	0.225	-0.641
	(-0.26)	(-0.88)	(0.47)	(-1.59)	(-0.23)	(0.65)	(-1.95)
Mean_mean_fl	2.640	2.575	0.325	4.355	5.230*	-0.169	0.415
	(0.71)	(0.85)	(0.11)	(1.86)	(2.49)	(-0.08)	(0.24)
Var_mean_fl	-23.34	-16.21	-1.674	4.040	-0.471	2.993	6.643
	(-1.24)	(-1.04)	(-0.11)	(0.32)	(-0.05)	(0.27)	(0.69)
Mean_var_fl	5.123	4.134	4.654*	1.878	4.605**	5.618***	2.000
	(1.88)	(1.88)	(2.14)	(1.09)	(3.02)	(3.73)	(1.46)
Var_var_fl	-15.15	2.427	15.90*	-0.202	4.188	-9.443	-6.865
	(-1.62)	(0.32)	(2.24)	(-0.03)	(0.88)	(-1.70)	(-1.53)
Mean_ep_y1	1.085	-1.228	1.673	0.525	0.713	0.373	0.346
	(0.57)	(-0.79)	(1.14)	(0.43)	(0.69)	(0.36)	(0.39)
Mean_mean_ep_yn	-2.211	1.657	-0.690	2.835	-1.217	-1.419	-1.068
	(-0.71)	(0.54)	(-0.21)	(0.98)	(-0.45)	(-0.47)	(-0.39)
Var_mean_ep_yn	10.80	6.349	-9.467	-18.87	-31.54*	-8.376	27.23*
	(0.64)	(0.43)	(-0.64)	(-1.35)	(-2.35)	(-0.61)	(2.15)
Constant	0.717	-1.347	-0.762	-4.411	-4.533	-1.898	1.576
	(0.12)	(-0.26)	(-0.15)	(-0.99)	(-1.14)	(-0.45)	(0.42)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 12. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Panel-Data Counterfactual Method using Corrected Akaike Information Criterion in Model (2) Generated Data with Treatment Period at 10

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_factor_y1	0.312 (1.54)	-0.539* (-2.35)	-0.236 (-1.22)	0.752** (2.66)	-0.118 (-0.50)	57.14 (0.38)	5661.1 (0.97)
Var_factor_y1	3.337*** (21.22)	3.730*** (20.35)	2.868*** (18.37)	3.801*** (17.23)	3.207*** (17.43)	30.15 (0.26)	6004.3 (1.31)
Mean_var_factor_yn	-0.681* (-2.39)	-0.823* (-2.26)	-0.488 (-1.36)	-0.568 (-1.06)	-1.477** (-3.11)	284.2 (0.83)	12678.0 (0.94)
Mean_mean_factor_yn	0.172 (0.51)	-0.291 (-0.63)	-0.403 (-0.92)	-0.165 (-0.24)	0.942 (1.59)	-63.65 (-0.15)	-5566.2 (-0.35)
Constant_y1	0.0523 (0.54)	0.109 (0.97)	-0.0266 (-0.28)	0.192 (1.41)	0.118 (1.02)	-144.5 (-1.93)	-2686.0 (-1.00)
Mean_constant_yn	0.0613 (0.36)	0.254 (1.11)	-0.293 (-1.34)	0.130 (0.38)	-0.395 (-1.30)	174.6 (0.83)	-9882.7 (-1.22)
Var_constant_yn	-0.212 (-1.50)	-0.191 (-1.01)	-0.161 (-0.95)	-0.478 (-1.79)	-0.00278 (-0.01)	37.17 (0.22)	-2543.4 (-0.41)
Mean_mean_fl	-0.367 (-0.30)	-0.116 (-0.08)	-1.327 (-1.06)	1.051 (0.58)	-0.594 (-0.42)	383.4 (0.43)	-18334.2 (-0.52)
Var_mean_fl	-0.716 (-0.11)	-0.444 (-0.06)	-11.96 (-1.76)	2.727 (0.29)	-0.429 (-0.06)	-2440.4 (-0.53)	-22136.8 (-0.13)
Mean_var_fl	5.572*** (5.87)	0.681 (0.66)	1.892* (2.09)	2.356 (1.82)	1.496 (1.42)	-144.1 (-0.22)	-9555.3 (-0.38)
Var_var_fl	-2.617 (-0.84)	-2.018 (-0.55)	-2.597 (-0.85)	-6.932 (-1.49)	-6.856* (-2.04)	1449.9 (0.58)	90122.5 (0.96)
Mean_ep_y1	1.322* (2.18)	0.0268 (0.04)	-0.403 (-0.66)	0.555 (0.64)	-0.295 (-0.42)	-174.6 (-0.38)	7378.1 (0.42)
Mean_mean_ep_yn	0.00262 (0.00)	-1.265 (-0.86)	0.490 (0.34)	1.509 (0.70)	-0.0298 (-0.02)	460.8 (0.35)	66007.9 (1.27)
Var_mean_ep_yn	-7.614 (-1.31)	-7.409 (-1.05)	4.584 (0.66)	10.45 (1.02)	-8.319 (-0.94)	-4300.5 (-0.64)	62274.3 (0.26)
Constant	-3.106 (-1.57)	3.481 (1.43)	3.102 (1.38)	-2.865 (-0.92)	3.751 (1.39)	-615.1 (-0.33)	-42698.6 (-0.60)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 13. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Panel-Data Counterfactual Method using Corrected Akaike Information Criterion in Model (2) Generated Data with Treatment Period at 20

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_factor_y1	-0.329* (-2.28)	-0.218 (-1.54)	-0.171 (-1.30)	0.116 (0.84)	-0.0910 (-0.85)	-0.0594 (-0.50)	0.0102 (0.10)
Var_factor_y1	2.673*** (21.81)	2.120*** (18.79)	2.559*** (24.52)	1.948*** (17.84)	1.555*** (18.72)	1.480*** (16.68)	1.463*** (17.11)
Mean_var_factor_yn	-0.366 (-1.83)	-0.814*** (-3.67)	-1.017*** (-4.21)	-1.047*** (-3.69)	-1.202*** (-5.35)	-1.296*** (-4.74)	-1.241*** (-5.03)
Mean_mean_factor_yn	0.229 (0.90)	-0.0355 (-0.13)	-0.359 (-1.22)	-0.0823 (-0.24)	-0.0253 (-0.09)	1.009** (3.03)	-0.0814 (-0.27)
Constant_y1	-0.0931 (-1.31)	-0.0407 (-0.61)	0.0581 (0.90)	0.0596 (0.84)	0.0515 (0.97)	-0.114 (-1.91)	0.0194 (0.37)
Mean_constant_yn	0.0316 (0.26)	-0.251 (-1.88)	-0.0769 (-0.51)	0.145 (0.87)	-0.000522 (-0.00)	-0.407* (-2.48)	0.119 (0.77)
Var_constant_yn	-0.0252 (-0.25)	0.0990 (0.92)	0.0174 (0.15)	0.164 (1.28)	-0.241* (-2.10)	0.0304 (0.25)	-0.0628 (-0.57)
Mean_mean_fl	-1.612 (-1.76)	-0.775 (-0.94)	-0.294 (-0.35)	0.800 (0.93)	0.0887 (0.13)	0.0225 (0.03)	0.729 (1.07)
Var_mean_fl	-1.467 (-0.32)	0.233 (0.05)	5.060 (1.22)	-5.746 (-1.18)	5.362 (1.57)	-2.080 (-0.54)	-0.337 (-0.10)
Mean_var_fl	0.533 (0.79)	1.961** (3.03)	1.675** (2.83)	0.769 (1.20)	0.660 (1.32)	0.822 (1.53)	-0.942* (-1.96)
Var_var_fl	4.273 (1.84)	-0.985 (-0.44)	-0.103 (-0.05)	-0.278 (-0.12)	2.832 (1.50)	3.351 (1.81)	3.321 (1.81)
Mean_ep_y1	0.321 (0.70)	-0.0650 (-0.16)	-0.666 (-1.59)	-0.477 (-1.10)	-0.109 (-0.31)	-0.479 (-1.26)	0.265 (0.78)
Mean_mean_ep_yn	-0.686 (-0.87)	-0.948 (-1.12)	-0.657 (-0.73)	0.141 (0.14)	-0.420 (-0.46)	0.291 (0.29)	-1.942 (-1.95)
Var_mean_ep_yn	-8.014 (-1.83)	-5.652 (-1.30)	-3.167 (-0.67)	5.003 (0.97)	-7.114 (-1.63)	1.266 (0.26)	0.594 (0.13)
Constant	3.832* (2.56)	2.884 (1.95)	2.370 (1.62)	1.227 (0.75)	2.855* (2.21)	2.648 (1.86)	4.587*** (3.38)

* p<0.05, **p<0.01, ***p<0.005.

[1] The regressand is the MSD of 1000 trials in each model setting.

Table 14. Coefficients (T-values) of Regression on Mean Square Deviation (MSD)[1] by the Panel-Data Counterfactual Method using Corrected Akaike Information Criterion in Model (2) Generated Data with Treatment Period at 30

	Maximum Region Number						
Regressors	4	5	6	7	8	9	10
Mean_factor_y1	4.791*** (7.54)	-1.863* (-2.45)	1.362* (2.53)	1.712*** (3.74)	-3.009*** (-4.18)	1.799*** (4.33)	2.139*** (7.52)
Var_factor_y1	2.115*** (12.37)	4.556*** (26.79)	2.306*** (19.74)	2.006*** (14.95)	3.931*** (21.58)	0.901*** (8.09)	0.585*** (7.57)
Mean_var_factor_yn	-0.159 (-0.64)	0.521 (1.29)	0.706* (2.15)	0.308 (1.27)	0.443 (1.15)	-0.0335 (-0.12)	0.00321 (0.02)
Mean_mean_factor_yn	-3.276** (-3.16)	-4.184** (-2.69)	-4.291** (-3.30)	-3.232** (-3.21)	-6.049*** (-3.52)	-2.037 (-1.87)	-1.689* (-2.06)
Constant_y1	0.162 (0.57)	-0.0976 (-0.26)	-0.0289 (-0.10)	-0.107 (-0.56)	0.0337 (0.10)	-0.559** (-3.14)	0.0638 (0.50)
Mean_constant_yn	-0.0953 (-0.19)	0.785 (1.03)	1.190* (2.00)	-0.466 (-0.96)	-0.586 (-0.72)	-0.401 (-0.79)	0.690 (1.79)
Var_constant_yn	-0.555 (-1.24)	0.579 (0.93)	-0.337 (-0.65)	-0.285 (-0.75)	-0.0602 (-0.09)	0.200 (0.49)	-0.464 (-1.60)
Mean_mean_fl	7.709* (2.11)	-3.032 (-0.64)	-6.003 (-1.71)	-0.284 (-0.11)	4.004 (0.97)	0.541 (0.23)	-0.377 (-0.23)
Var_mean_fl	23.33 (1.22)	-18.43 (-0.74)	-10.62 (-0.59)	-1.915 (-0.14)	11.67 (0.54)	-12.60 (-1.00)	15.41 (1.65)
Mean_var_fl	9.469*** (3.46)	0.127 (0.04)	4.235 (1.61)	0.238 (0.13)	3.872 (1.27)	0.681 (0.41)	-0.667 (-0.56)
Var_var_fl	-0.740 (-0.08)	-20.99 (-1.74)	-2.030 (-0.23)	-2.004 (-0.31)	-13.68 (-1.28)	2.146 (0.37)	2.192 (0.55)
Mean_ep_y1	0.113 (0.06)	0.554 (0.22)	-1.799 (-1.06)	-0.255 (-0.21)	0.483 (0.23)	-0.161 (-0.14)	0.0439 (0.05)
Mean_mean_ep_yn	3.288 (1.08)	-6.790 (-1.41)	-0.0943 (-0.02)	-0.483 (-0.16)	-5.737 (-1.08)	3.221 (1.03)	3.103 (1.26)
Var_mean_ep_yn	1.622 (0.10)	-38.94 (-1.63)	59.78** (3.08)	-3.162 (-0.21)	8.493 (0.31)	-2.696 (-0.18)	9.598 (0.83)
Constant	-14.42** (-3.09)	9.189 (1.54)	5.254 (1.15)	3.631 (1.08)	-0.221 (-0.04)	1.411 (0.47)	2.984 (1.39)

* $p<0.05$, ** $p<0.01$, *** $p<0.005$.

[1] The regressand is the MSD of 1000 trials in each model setting.