Hindawi
Journal of Mathematics
Volume 2022, Article ID 6211454, 8 pages
https://doi.org/10.1155/2022/6211454

Research Article

Principal Covariates Regression for Causal Case Studies

Weibin Ding,[1] Jie Li,[2] Dian Jin,[3,4] and Jiayang Kong[3,4,5]

[1]*State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310012, China*
[2]*Jinhua Power Supply Company of State Grid Zhejiang Electric Power co., Ltd., Jinhua 321000, China*
[3]*School of Economics and Business Administration, Central China Normal University, Wuhan 430079, China*
[4]*School of Mathematics and Statistics, Central China Normal University, Wuhan 430079, China*
[5]*Research Center of Low-Carbon Economy and Environmental Policies, Central China Normal University, Wuhan 430079, China*

Correspondence should be addressed to Jiayang Kong; jiayangkong@hotmail.com

Received 15 June 2022; Accepted 20 September 2022; Published 15 October 2022

Academic Editor: Niansheng Tang

Researcher and analyst are often interested in estimating the effect of an intervention or treatment, which takes place at the aggregate level and affect one single unit, such as country and region. Thus, comparative case studies would be their first choice in practice. However, comparative case studies could fail to yield an estimate in the effect that is unbiased and consistent, as in some contexts; there are not suitable control units that are similar to the treated. The econometric literature has taken synthetic control methods and panel data approaches to this problem. In this study, we developed a principal covariate regression estimator, which exploits the cross-sectional correlation, as well as the temporal dependency, to reproduce the dynamics of the treated in the absence of an event or policy. From a theoretical perspective, we introduce the statistical literature on dimensional reduction to make a causal inference. From a technique perspective, we combine the vertical regression and the horizontal regression. We constructed an annual panel of 38 states, to evaluate the effect of Proposition 99 on beer sales in California, using the principal covariate regression estimator proposed here. We find that California's tobacco control program had a significant negative and robust effect on local beer consumption, suggesting that policymakers could reduce the use of cigarette and alcohol in the public using one common behavioral intervention.

1. Introduction

Randomized controlled trials are considered as the gold standard for scientific studies, which compare the outcome variable in the treatment group with that in the control group [1–3]. They yield an unbiased and consistent estimate for the effect of the intervention or treatment, and provide the strongest level of evidence on the causal interpretation [4, 5]. Randomized controlled trials, however, have their drawbacks [6–8]. For example, they are a time-consuming task, and cost a lot of money [9]. Besides, some studies are not likely to be ethically done using randomized controlled trials [10].

Instead, researcher and analyst often perform a comparative case study to evaluate the effect of an event or policy, particularly when the intervention or treatment happens to one single unit at an aggregate level, such as country and region [11, 12]. In the comparative case study, research and analyst compare the dynamics of the aggregate outcome for the treated unit to those for a set of controls that are not affected by the event or policy, and obtain an estimate in the average treatment effect in the treated over the period after the introduction of an intervention or treatment [13–17]. For example, Card and Kruger measured the effectiveness of the minimum wage on the unemployment, by comparing New Jersey with Pennsylvania, based on data on fast food restaurants [18].

However, comparative case studies have some limitations, which would damage their credibility of the causal relationship between intervention and outcome, as well as limit their application in practice [19]. For example, it is subject for research and analyst to choose one unit without experiencing an event or policy, which is used as the counterfactual of the united [20]. Also, in certain contexts, it is possible that research and analyst cannot find one control that is suitable to be the counterfactual for the unit affected

[21]. Besides, as research and analyst often use a small sample size in comparative case studies, the additive technique is needed to interpret the significance of the effect of an intervention or treatment [22].

In this paper, we address these challenges. First, we develop a novel approach to estimate the average treatment effect in comparative case studies, which is a data-driven approach that forms a comparative unit for the treated. The literature on econometrics has advocated two alternative approaches for program evaluation in the setting with only one unit receiving an intervention or treatment [23, 24]. The synthetic control method constructs the counterfactual for the treated using a combination of similar units that are not affected by an event or policy [25]. The panel data approach exploits the cross-sectional correlation to form one comparative unit for the treated [26]. Although at first sight different, the two approaches are methodologically quite similar in terms of the pattern in the data that they adopt to reproduce the dynamics of the outcome variable in the treated in the absence of an intervention or treatment [27]. Here, we introduce principal covariate regression from statistical literature on dimensional reduction to make a causal inference for comparative case studies. Principal covariate regression is an approach, where we regress a collection of outcome variables with regard to a collection of characteristic variables, particularly when the number of the latter is large or it is collinear among them. Principal covariate regression also is referred to as a technique to select characteristic variable [28].

Second, we propose the bootstrapping test to interpret the significance of the effect. That is, we first reconstruct a new control group by the bootstrapping technique, and then estimate the effect of an event or policy based on the bootstrapping sample. We repeat the process above 1000 times, and obtain an empirical distribution of the effect. Using the empirical distribution, we could calculate the empirical standard error for the effect, and make an interpretation of the effective significance. The bootstrapping test proposed here could be used whether the data are at the individual level or at the aggregate level, and it does not require that there are a large number of units in the control group.

Finally, as an illustrative example, we estimated the effect of Proposition 99 on beer sales. There exists wide agreement among epidemiologists that smoking and drinking have a positive association, although the reason why they are associated remains unclear28. Estimating the effect of this policy requires that we produce the counterfactual for California; that is, the situation of California in the absence of Proposition 99. We used the principal covariate regression estimator developed in this study to reproduce the dynamics of California by exploiting the cross-sectional correlation and the temporal dependency. Estimated results show Proposition 99 had a negative effect on beer sales in California. This result also has an important implication for policy-relevant questions, suggesting that policymakers could reduce the sales in smoking and alcohol using one common intervention.

The rest of this paper is organized as follows: Section 2 explains the major ideas behind principal covariate regression for causal comparative case studies. In Section 3, we used the principal covariate regression estimator to assess the effect of California's tobacco control program on local beer consumption. Section 4 concludes.

2. Methods

In comparative case studies, researcher and analyst often compare one unit affected by an event or policy with other units that are not affected. These unaffected units also are called the control units in the literature related to econometrics, which are referred to as the counterfactual of the treated unit. Therefore, it is quite important for researcher and analyst to find one suitable comparison unit that is similar to the treated unit before the introduction of an intervention or treatment in comparative case studies, which directly affects the credibility of the result. In this section, we propose a new approach, which is drawn on the literature on econometrics and computer science, to construct the counterfactual for the treated unit.

2.1. A Motivating Model. In this subsection, we explain the idea behind our new approach. Suppose that we have $J+1$ units. Without loss of generality, also suppose that the first unit experiences an event or policy at certain time. Thus, we have J control units, that is, the J remaining units.

Let $Y_{it}^{\text{unaffected}}$ be the potential outcome, which would be observed for unit i at time t if unit i is unexposed during the period before the introduction of an intervention or treatment. Let Y_{it}^{affected} be the potential outcome, which would be observed for unit i at time t if unit i is exposed during the period after the introduction of an intervention or treatment. Let Y_{it}^{observed} be the actual outcome observed for unit i at time t in the comparative case study. Therefore, we have $Y_{it}^{\text{unaffected}} = Y_{it}^{\text{observed}}$ for unit i at time t during the period before the introduction of an intervention or treatment. We also assume that an intervention or treatment takes place at time T_0.

Based on these definitions above, we could obtain the dynamic effects of an event or policy on an outcome of our interest in the comparative case study using the following equation:

$$\alpha_{1t} = Y_{1t}^{\text{affected}} - Y_{1t}^{\text{unaffected}}, \text{for } t \geq T_0, \tag{1}$$

where α_{1t} denotes the effect of an event or policy for the first unit at time t, during the period after the introduction of an event or policy.

Because $Y_{1t}^{\text{affected}} = Y_{1t}^{\text{observed}}$ for the first unit at time t during the period after the introduction of an intervention or treatment, in order to obtain α_{1t}, we need to estimate $Y_{1t}^{\text{unaffected}}$ for the first unit at time t during the period after the introduction of an intervention or treatment.

The econometric literature has advocated two approaches to estimate $Y_{1t}^{\text{unaffected}}$ [23, 24]. Abadie et al. advised to use a weighted average of control units as the estimate of $Y_{1t}^{\text{unaffected}}$ [25], that is,

$$Y_{1t}^{\text{unaffected}} = \sum_{j=2}^{J+1} w_j Y_{jt}^{\text{unaffected}} = \sum_{j=2}^{J+1} w_j Y_{jt}^{\text{observed}}, \quad (2)$$

where w_j represents the weight, which is obtained by minimizing the distance of characteristic variables between the first unit and control units.

On the other hand, Hsiao and Zhou recommended researcher and analyst used a linear regression, which exploits the cross-sectional correlation between the first unit and control units, to estimate $Y_{1t}^{\text{unaffected}}$ [26]. The linear regression fits the following model:

$$Y_{1t}^{\text{unaffected}} = \alpha + \sum_{j=2}^{J+1} \beta_j Y_{jt}^{\text{unaffected}} + \varepsilon_{1t} = \alpha + \sum_{j=2}^{J+1} \beta Y_{jt}^{\text{observed}} + \varepsilon_{1t}, \quad (3)$$

where α is the constant item. β_j represents the regressive coefficient. ε_{1t} is the error item.

Equations (1) and (2) also are written using the following function:

$$\begin{aligned} Y_{1t}^{\text{unaffected}} &= f\left(Y_{2t}^{\text{unaffected}}, Y_{3t}^{\text{unaffected}}, \cdots, Y_{J+1,t}^{\text{unaffected}}\right) \\ &= f\left(Y_{2t}^{\text{observed}}, Y_{3t}^{\text{observed}}, \cdots, Y_{J+1,t}^{\text{observed}}\right). \end{aligned} \quad (4)$$

We could fit the above function using some standard estimation methods, although the function form $f(\cdot)$ is unknown. In this paper, we introduce principal covariate regression from the statistical literature on dimensional reduction to estimate the function form above, which imputes the missing status for the first unit in the absence of an intervention or treatment. Next, we would explain how principal covariate regression works in details.

2.2. A Brief Introduction to Principal Covariate Regression. In this subsection, we briefly introduce principal covariate regression. Suppose that Y and X be two matrices, of order $N \times K$ and $N \times J$, respectively, where N represents the number of observations, and K and J denote the number of outcome variables and the number of characteristic variables, respectively. In principal component regression, R components could be expressed as a weighted combination of the matrix of X using the following formula:

$$T = XW, \quad (5)$$

where T is a $N \times R$ matrix of the component score, and W is a $J \times R$ matrix of the component weight. The component score plays a role in explaining both of Y and X, namely.

$$Y = TP_Y + E_Y, \quad (6)$$

where P_Y is the $R \times K$ matrix of the weights for the K outcome variables on the R components.

Using the Equation (6), we obtain predictions of outcome variables:

$$\widehat{Y} = T\widehat{P}_Y, \quad (7)$$

where $\widehat{P}_Y$ is the estimates in P_Y, and $\widehat{Y}$ is the prediction of Y.

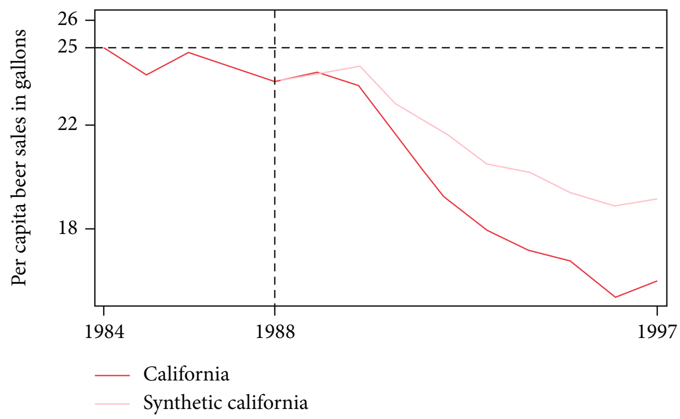
So far, we briefly introduce principal covariate regression. Next, we illustrate how to use the principal covariate regression estimator to estimate the effect of an event or policy.

2.3. Implementation of the Principal Covariate Regression Estimator. In the previous subsection, we described the principal covariate regression. Now, we illustrate how we use the principal covariate regression estimator to estimate the effect of an event or policy in comparative case studies. First, we trained the principal covariate regression model based on the data on all the units before an event or policy with a lasso technique. In this step, we referred to values of the treated unit as outcome variable, while we referred to values of the control group as characteristic variables. We then used the trained principal covariate regression model to impute the missing status for the treated unit after the introduction of an intervention or treatment. Finally, we compared the observed values with the imputing values, and obtained the dynamic effects of an event or policy.

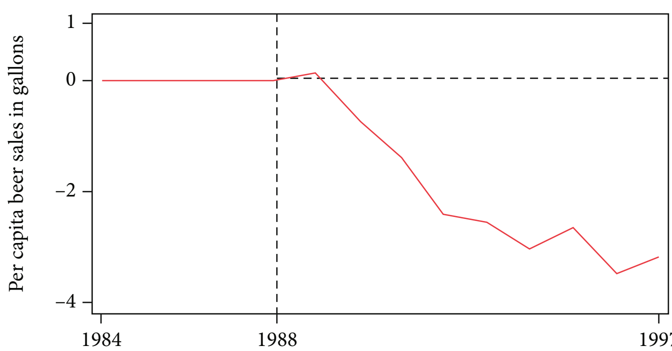
2.4. Inference in Comparative Case Studies. Large sample inference is not suitable for comparative case studies, as there are a small number of units included in them. Abadie et al. advised researcher and analyst performing a series of placebo studies to interpret the significance of the results. Previous research, however, pointed out that the results from the placebo study could be distorted due to the size of control units. Here, we proposed the bootstrapping test for comparative case studies, which reconstructs the control group and generates an empirical distribution of the effect that an event or policy had on an outcome.

3. Estimating the Effect of Proposition 99 on Beer Sales

3.1. Background. As public awareness of health risk of smoking had increased dramatically over the past decade, the government of California launched a tobacco control program in the year of 1989, that is, Proposition 99, leading a new wave of antitobacco legislation at state and federal levels across the United States. Proposition 99 aims to reduce the behavior of smoking by raising the tax on cigarette. Previous research has reported that Proposition 99 had a statically significant and negative effect on cigarette sales in California. Yet, the evidence on whether Proposition 99 affected local beer sales is still unknown. Considering that smoking and drinking always go together, not only is an investigation into the effect of tobacco control program on beer consumption quite interesting but it also has an important implication for policy-relevant questions. That is, if we find a negative



(a) California and synthetic California



(b) The gap in beer consumption between two states

FIGURE 1: The effect of California's Proposition 99 on local beer consumption.

association between California's Proposition 99 and local beer sales, then it allows policymakers to decrease the use of cigarette and alcohol using one common behavioral intervention.

3.2. Data and Sample. In order to assess the effect of Proposition 99 on beer sales, we constructed an annual panel including 38 states in the United States, from 1984 to 1997. We excluded other states because these states passed an analogous tobacco control program during the period 1989 to 1997. As California's tobacco control program was launched in 1989, our study period contains 5 years before the introduction of Proposition 99, and 9 years after that. The outcome variable of our interest is annual per capita beer consumption at state level, which is measured in our data as per capita sales in gallons. Using the data and the technique that is described in the previous section, we created the counterfactual for California. Following the econometric literature, we name California's counterfactual the synthetic California. We estimated the dynamic effects of Proposition 99 on beer sales as the differences in levels of per capita beer consumption between California and its counterfactual, that is, synthetic California, in the years after the introduction of Proposition 99.

3.3. Results. Panel A in Figure 1 plots the dynamics of per capita beer consumption in California and the synthetic California, where the dark red line represents California

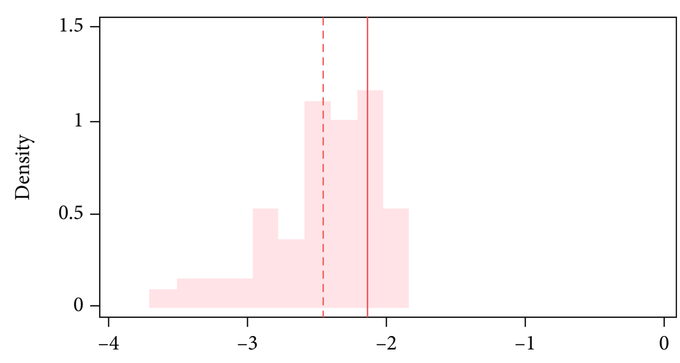
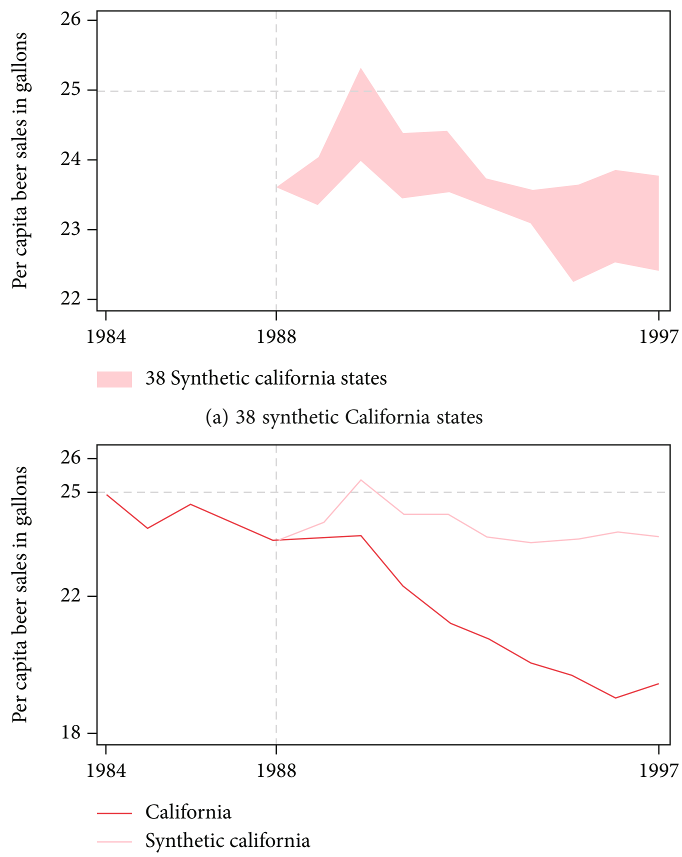


FIGURE 2: The results from the bootstrapping test.

and the light red line represents the synthetic California. As this panel shows, the synthetic California provides a comparative unit suitable for California to assess the effect of Proposition 99 on per capita beer sales. That is, before the government of California launched the tobacco control program, the time path of beer consumption in California was almost perfectly overlaid with that in the synthetic California. From this panel, we also see that levels of beer consumption still were similar in California and the synthetic California in the late 1980s. They, however, began to diverge in the early 1990s, when Proposition 99 had been



(a) 38 synthetic California states

(b) California and 38 synthetic California states

FIGURE 3: The results from the leave-one-state-out test.

passed for about two years. The lag effect observed could be because California's tobacco control program first reduces the use of cigarette, which gradually promotes public health behaviors, and consequently leads to the use of alcohol declines. In other words, the reduction in beer consumption might be one side effect of California's tobacco control program, which responses to this program over a relatively long term. In addition, this panel shows that beer sales continued to decline after the introduction of Proposition 99, but with a larger decrease in California than in the synthetic California. In the year of 1996, the difference in beer consumption between the two states arrived at the maximum. That is, per capita beer consumption was about 4 points higher in the synthetic California relative to California.

Panel B in Figure 1, plotting the dynamics of the differences in per capita beer sales between California and the synthetic California, assures our findings. As this panel shows, California's Proposition 99 had a persistent effect on local beer consumption. That is, the gap had been widening over the years after the introduction of Proposition 99, even if there was a reduction of gap in 1997, the last year in our study period.

3.4. Inference about the Effect of California's Proposition 99 on Local Beer Consumption. In order to interpret the significance of the estimated results, we conducted the bootstrapping test. That is, first, we drew, with replacement, 38 states, with entire observations between 1984 and 1997 for each state, from our original sample, which forms a new sample. We name the new sample the bootstrapping sample. We then estimated the effect of California's tobacco control program on per capita beer sales based on the bootstrapping sample. The above process is repeated 1000 times, and provides an empirical distribution of the effects that Proposition 99 had on beer consumption.

Figure 2 plots the results from the bootstrapping test. The red shadow represents the empirical distribution of the effect that California's tobacco control program on per capita beer sales. The dash red line is the mean of the empirical distribution. The solid red line is the estimated effect of Proposition 99 obtained from the original sample. As this figure shows, the empirical distribution of the effect that California's tobacco control program had on local beer consumption is far away from the zero, while the two vertical lines are close to each other, suggesting that Proposition 99 had a statistically significant and substantially negative effect on per capita beer sales in California.

3.5. Sensitivity Analysis. As we said in the previous, the major problem of comparative case studies is the subject choice of comparative units by researcher and analyst, which might harm the credibility of the estimation from comparative case studies. Although a data-driven approach could alleviate the concern, the observed results probably motivated by a particular comparative unit exactly existing in the original sample. Therefore, in order to increase our confidence that Proposition 99 had a substantial negative effect on beer consumption in California, as well as to interpret that our results are not driven by a particular control state included in our sample, we performed the leave-one-state-out test by iteratively applying the principal covariate regression estimator to construct 38 synthetic California states based on 38 subsamples, with each sample excluding one state from the original sample in turn.

Figure 3 presents the results from the leave-one-state-out test. In Panel A, we plot the 38 synthetic California states. As this panel shows, all the time paths of per capita beer sales for the 38 synthetic California states are close to each other, indicating that the creation of the synthetic California is not sensitive to the choice of control states that are used to form the synthetic California. In Panel B, we add the dynamics of California into Panel A. As this panel shows, all the 38 synthetic California states have a different trend in per capita beer sales from California after 1990, the second year after the introduction of Proposition 99. This observation, from the Panel B, further increases our confidence that California's tobacco control program had a lag and substantial negative effect on local beer consumption.

4. Conclusion

When there is merely one unit experiencing an event or policy, in order to assess the effect of the event or policy, the first and the desired choice for researcher and analyst is to do a comparative case study. However, in practice, the implementation of comparative case studies has limitations. On one hand, it is difficult for researcher and analyst to find a suitable comparative unit. On the other hand, the choice of control units used to form a comparative unit suitable for the treated unit depends on individual experience of researcher and analyst. That is, different people would hold various choices, which probably leads to distinguished estimated results of the effect of the same intervention or treatment. This heavily reduces the credibility of estimated results from comparative case studies.

In this paper, we developed a new approach for estimating the average causal effect in a comparative case study, that is, the principal covariate regression estimator. This new estimator avoids the subject choice of control units by researcher and analyst, and adopts a data-driven procedure to create a comparative unit suitable for the treated unit, which reproduced the dynamics of the outcome variable for the treated unit before an event or policy. We also explained the implementation of the long short-term estimator in details. Besides, we proposed the bootstrapping test to interpret the significance of the estimated results in a small sample, particularly in comparative case studies, where the larger sample inferential technique is often not suitable.

As an illustrative example, we estimated the effect of California's tobacco control program on local beer consumption. Epidemiological research reported that cigarette and alcohol always went together, even though the exact nature of this association between them is still little known. Previous studies also evaluated the effect of Proposition 99 on cigarette consumption in California. However, the evidence on the effect of Proposition 99 on California's beer sales remains unclear. Our results provide evidence that California's tobacco control program had a substantial and persistent negative effect on local beer consumption. In the bootstrapping test, we also observed a statistically significant negative Proposition 99's effect on per capita beer sales. In addition, using the leave-one-state-out test, we interpret that our construction of the synthetic California are not sensitive to the choice of control units, again.

Our study contributes to the literature on causal inference in comparative case studies. We provide a novel approach to construct the counterfactual that is comparative to the treated unit, which adopts the principal covariate regression to predict the potential outcome for the treated unit. On the other hand, our work expand the coverage of the application of machine learning. To our knowledge, it is the first to apply the principal covariate regression to causal inference in comparative case studies.

Of course, our study has limitations. In this paper, we only consider a special setting, where just one unit is exposed to the event or intervention. In real word, there might be a set of units exposed to at event or intervention. Our approach proposed here could not be directly applied to these settings. However, we could use control units to construct the counterfactual for each treated unit based on the regression tree model, and estimate the individual effect for each treated unit. Besides, our approach heavily depends on the information on the correlation across units, which does not consider the time dependency of outcome of the treated unit. Future research should try to construct the counterfactual of the treated, simultaneously using them. For the potential application, the approach can also be used in the energy field, for example, to analyze the influence of the design of subsidy policies for the new energy industry on installed capacity. We are using this approach to study the impact of regulatory policy changes on grid cost monitoring and audit costs. Preliminary results show that regulatory policies significantly affect cost monitoring and audit costs. Additionally, the approach proposed in this paper also is applied in the fields of economics and management, where often there exists many issues related to causal inference. For example, economists often want to understand the effect of job train program on income. Thus, our approach would had a larger range of the potential application.

Data Availability

Data are available from corresponding author upon reasonable request.

Additional Points

Correspondence and requests for materials should be addressed to Jiayang Kong.

Disclosure

This paper does not reflect an official statement or opinion from the organizations.

Conflicts of Interest

The authors declare no competing interests.

Authors' Contributions

All authors equally contributed to this paper. Weibing Ding and Jie Li conceptualized the study and provided supervision. Dian Jin and Jiayang Kong collected the data, conducted the statistical analysis, and drafted the manuscript. Weibing Ding and Jie Li contributed to the interpretation of the results. All authors provided critical feedback on drafts and approved the final manuscript.

Acknowledgments

We acknowledge the funding from the Science and Technology Project of State Grid Corporation of China (Grant no. B311JH21000D). We thank Lang Cheng, Wei Wei, and Yihui Li for their helpful comments on an earlier version of this paper. In addition, Jiayang Kong wants to thank, in particular, for the patience, care, and support from Xiaoying Zhou over the years.

References

[1] M. L. Meldrum, "A brief history of the randomized controlled trial: from oranges and lemons to the gold standard," *Hematology/Oncology Clinics of North America*, vol. 14, no. 4, pp. 745–760, 2000.

[2] E. Hariton and J. J. Locascio, "Randomised controlled trials – the gold standard for effectiveness research," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 125, no. 13, pp. 1716–1716, 2018.

[3] S. Webber and C. Prouse, "The new gold standard: the rise of randomized control trials and experimental development," *Economic Geography*, vol. 94, no. 2, pp. 166–187, 2018.

[4] D. T. Wade, "Randomized controlled trials – a gold standard?," *Clinical Rehabilitation*, vol. 13, no. 6, pp. 453–455, 1999.

[5] N. M. Castillo and D. A. Wagner, "Gold standard? The use of randomized controlled trials for international educational policy," *Comparative Education Review*, vol. 58, no. 1, pp. 166–173, 2014.

[6] T. J. Kaptchuk, "The double-blind, randomized, placebo-controlled trial: gold standard or golden calf?," *Journal of Clinical Epidemiology*, vol. 54, no. 6, pp. 541–549, 2001.

[7] N. Cartwright, "Are RCTs the gold standard?," *BioSocieties*, vol. 2, no. 1, pp. 11–20, 2007.

[8] I. H. Kvangraven, "Impoverished economics? A critical assessment of the new gold standard," *World Development*, vol. 127, article 104813, 2020.

[9] V. Tyagi and S. Webber, "A rusting gold standard: failures in an Indonesian RCT, and the implications for poverty reduction," *Environment and Planning A: Economy and Space*, vol. 53, pp. 992–1011, 2021.

[10] G. Thomas, "After the gold rush: questioning the "Gold standard" and reappraising the status of experiment and randomized controlled trials in education," *Harvard Educational Review*, vol. 86, no. 3, pp. 390–411, 2016.

[11] N. Brunner, A. Lele, M. Starkl, and L. Grassini, "Water sector reform policy of India: experiences from case studies in Maharashtra," *Journal of Policy Modeling*, vol. 32, no. 4, pp. 544–561, 2010.

[12] A. H. Samadi, S. Owjimehr, and Z. N. Halafi, "The cross-impact between financial markets, Covid-19 pandemic, and economic sanctions: the case of Iran," *Journal of Policy Modeling*, vol. 43, no. 1, pp. 34–55, 2021.

[13] D. Borbely, "A case study on Germany's aviation tax using the synthetic control approach," *Transportation Research Part A: Policy and Practice*, vol. 126, pp. 377–395, 2019.

[14] S. Hasancebi, "The economic cost of revolution: the Iranian case. A synthetic control analysis," *The Singapore Economic Review*, vol. 67, no. 1, pp. 267–287, 2022.

[15] V. Geloso and J. B. Pavlik, "The Cuban revolution and infant mortality: a synthetic control approach," *Explorations in Economic History*, vol. 80, article 101376, 2021.

[16] X. Zhang, G. Owen, M. Green, I. Buchan, and B. Barr, "Evaluation of the impact of tiered restrictions introduced in England during October and December, 2020, on COVID-19 cases: a synthetic control study," *The Lancet*, vol. 398, p. S92, 2021.

[17] F. España, R. Arriagada, O. Melo, and W. Foster, "Forest plantation subsidies: impact evaluation of the Chilean case," *Forest Policy and Economics*, vol. 137, article 102696, 2022.

[18] D. Card and A. B. Krueger, "Wages and employment: minimum a case study of the fast-food industry in New Jersey and Pennsylvania," *American Economic Review*, vol. 84, 1994.

[19] A. Abadie, A. Diamond, and J. Hainmueller, "Comparative politics and the synthetic control method," *American Journal of Political Science*, vol. 59, no. 2, pp. 495–510, 2015.

[20] S. Klößner, A. Kaul, G. Pfeifer, and M. Schieler, "Comparative politics and the synthetic control method revisited: a note on Abadie et al. (2015)," *Swiss Journal of Economics and Statistics*, vol. 154, no. 1, p. 11, 2018.

[21] S. Mahant and M. Hall, "Methodological progress note: interrupted time series," *Journal of Hospital Medicine*, vol. 16, no. 6, pp. 364–367, 2021.

[22] A. Abadie, "Using synthetic controls: feasibility, data requirements, and methodological aspects," *Journal of Economic Literature*, vol. 59, no. 2, pp. 391–425, 2021.

[23] A. Abadie and J. Gardeazabal, "The economic costs of conflict: a case study of the Basque Country," *American Economic Review*, vol. 93, no. 1, pp. 113–132, 2003.

[24] C. Hsiao, H. Steve Ching, and S. Ki Wan, "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*, vol. 27, no. 5, pp. 705–740, 2012.

[25] A. Abadie, A. Diamond, and J. Hainmueller, "Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, vol. 105, no. 490, pp. 493–505, 2010.

[26] C. Hsiao and Q. Zhou, "Panel parametric, semiparametric, and nonparametric construction of counterfactuals," *Journal of Applied Econometrics*, vol. 34, no. 4, pp. 463–481, 2019.

[27] S. Athey, M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi, "Matrix completion methods for causal panel data models," *Journal of the American Statistical Association*, vol. 116, no. 536, pp. 1716–1730, 2021.

[28] B. Jones and A. Artemiou, "On principal components regression with Hilbertian predictors," *Annals of the Institute of Statistical Mathematics*, vol. 72, no. 2, pp. 627–644, 2020.