Generalized Synthetic Control Method: Causal Inference with Instrumented Principal Component Analysis

Cong Wang*

February 6, 2024

Abstract

To address the limitations imposed by the parallel trend assumption (PTA) required by prevalent methods such as difference-in-differences (DID), synthetic control methods (SCM) leverage data from the control group to impute the missing counterfactual for the treated group post-treatment. However, the original SCM and its derivatives primarily rely on outcome data for these imputations, requiring that the outcomes of treated units fall within or close to the convex hull of the donor pool. Yet, in many instances, treated units are poorly represented by the donor pool. This paper expands the linear interactive fixed effects model by integrating covariates into dynamic factor loadings that interact with time-varying factors. This methodology confers multiple benefits: firstly, it incorporates the strengths of previous SCM approaches, such as the relaxation of the PTA and conditional randomization of treatment assignment. Secondly, it eliminates the need for correct functional form assumptions. Thirdly, by utilizing the dimension reduction capability of principal component analysis (PCA), it efficiently manages high-dimensional data, enhancing the value extracted from numerous covariates.

Keywords: Synthetic Control, Principal Component Analysis, Causal Inference

JEL Codes: G11, G12, G30

*Department of Economics and Law, Sapienza University of Rome

1 Introduction

In this paper, we propose a new counterfactual imputation method that leverage the dimension reduction capability of instrumented principal component analysis (IPCA) to enhance the value extracted from numerous covariates. We name the newly proposed method the generalized synthetic control with instrumented principal component analysis (GSC-IPCA) in line with the privious generalized synthetic control method with interactive fixed effects (GSC-IFE) proposed by Xu (2017). The GSC-IPCA estimator is designed to overcome the constraints of the PTA, a requirement for widely adopted methodologies such as DID. Furthermore, it addresses the limitation observed in the original SCM and its variants, which predominantly depend on outcome data to impute missing counterfactuals, thus not fully leveraging the available covariate information. The GSC-IPCA estimator extends the linear interactive fixed effects model by incorporating covariates through unit-time-varying factor loadings interacted with time-varying factors, instrumented for additional robustness[1].

Causal inference in economics and other social sciences is often complicated by the absence of a counterfactual, which is essential for evaluating the impact of a treatment or policy intervention. The counterfactual represents the outcome that would have been observed in the absence of the treatment. The most common approach to estimating the counterfactual is the DID, which compares the average change in the outcome variable for the treated group with the average change for the control group. However, DID requires the PTA, which posits that, in the absence of the treatment, the average outcomes for the treated and control groups would have followed the same trend. This assumption is often difficult to verify and may be violated in practice. In contrast, SCM utilize control group data to estimate the missing counterfactuals for the treated group after treatment, functioning akin to a vertical regression, thereby offering an alternative when DID's underlying assumptions are untenable.

[1]The functional form of the data generating process in Xu (2017) only sepcifies unit-varying factor loadings.

The SCM estimates missing counterfactuals for treated units by constructing a weighted average of outcomes from control units. These weights are selected to ensure that pre-treatment outcomes for the control group closely align with those of the treated units. Diverging from the original SCM and its derivatives, which rely predominantly on outcome data for imputing post-treatment counterfactuals without requiring a comprehensive estimation of the data-generating process (DGP), the GSC-IFE seeks to explicitly model the DGP. Our proposed method advances the GSC-IFE framework by introducing a distinctive strategy for integrating covariates, thereby enhancing the method's capability to generate more accurate counterfactual predictions.

2 Framework

Consider Y_{it} as the observed outcome for a specific unit i at time t. The total number of units is $N = N_{treat} + N_{\text{ctrl}}$, where N_{treat} indicates the number of units in the treatment group, and N_{ctrl} represents those in the control group. Each unit is observed over T time periods, ranging from period 1 to period T. Let T_{pre} denote the number of pre-treatment periods, and T_{post} the number of post-treatment periods. The unit is treated at time $T_{\text{pre}}+1$, and the treatment effect is initially observed at time $T_{\text{pre}}+1$ and continues to be observed thereafter, a scenario commonly referred to as staggered adoption.

Assumption 1 *Functional form:*

$$\begin{aligned} Y_{it} &= D_{it} \circ \delta_{it} + B_{it} F_t + \mu_{it}, \\ B_{it} &= X_{it}\Gamma + H_{it} \end{aligned} \tag{1}$$

where D_{it} is a binary treatment indicator and δ_{it} signifies the treatment effect, which exhibits variation across units and through times[2]. The expression $B_{it} = [\beta_{it}^1, \ldots, \beta_{it}^K]$ represents a vector of factor loadings (the number of common factors is K.), whereas $F_t = [f_t^1, \ldots, f_t^K]'$

[2]The symble "$\circ$" represents point-wise product.

corresponds to a vector of time-varying common factors, and μ_{it} is the idiosyncratic error term. A key distinction of the proposed model from that delineated in Xu (2017) is the incorporation of factor loadings B_{it}, which are instrumented by observed covariates X_{it}. This integration permits B_{it} to exhibit variability across time and units, thereby introducing an additional layer of heterogeneity into the model.

The vector $X_{it} = [x_{it}^1, \ldots, x_{it}^L]$ consists of observed covariates, where L denotes the number of covariates. The factor loadings B_{it} are theorized to be a linear function of these observed covariates X_{it}, with Γ acting as the $L \times K$ coefficient matrix, and $H_{it} = [\eta_{it}^1, \ldots, \eta_{it}^L]$ comprising the vector of error terms.

Upon examining the functional form presented in Equation 1, we can amalgamate the two segments to formulate the ensuing equation:

$$Y_{it} = D_{it} \circ \delta_{it} + (X_{it}\Gamma)F_t + \epsilon_{it}, \quad \epsilon_{it} = \mu_{it} + H_{it}F_t. \tag{2}$$

The factor component of the model, $B_{it}F_t = \beta_{it1}f_{1t} + \beta_{it2}f_{2t} + \cdots + \beta_{itk}f_{kt}$, where $B_{it} = X_{it}\Gamma$, is assumed to adopt a linear, additive form. Despite appearing to be restrictively structured, this approach is capable of capturing a vast array of unobserved heterogeneities. It is inclusive of all specifications present in the interactive fixed effects model within the GS-IFE, such as unit and time fixed effects, unit-specific linear or quadratic time trends, and autoregressive processes. Beyond the additive integration of the treatment effect as delineated in Equation 2, the model imposes no additional constraints on the functional form of the treatment effect. This level of flexibility enables the straightforward application of PCA for estimating the factor loadings and common factors, thereby facilitating the imputation of counterfactual outcomes for treated units.

The main quantity of interest of this paper is the average treatment effect (ATE) for the treated, which is defined as:

$$\widehat{ATT}_t = \frac{1}{N_{treat}} \sum_{it} \left(Y_{it}(1) - \hat{Y}_{it}(0) \right) = \frac{1}{N_{treat}} \sum_{it} \hat{\delta}_{it}. \quad for \quad \forall i > N_{co}, \forall t > T_{pre}. \tag{3}$$

2.1 Assumptions for identification

Assumption 2 *Unconfoundedness:*

$$\epsilon_{it} \perp D_{js}, X_{js}, F_s \quad \forall i, j, s, t. \tag{4}$$

Assumption 2 stipulates that the error term for any unit at any time period is independent of treatment assignment, observed covariates, and unobserved time-varying factors. This independence is a crucial condition that lends substance to model Equation 2 and is imperative for the consistent estimation of Γ.

Assumption 3 *Regularity conditions: (1)* Γ *is bounded and has a finite second moment, (2)* F_t *is bounded and has a finite second moment, (3)* X_{it} *is bounded and has a finite second moment.*

The regularity conditions outlined in Assumption 3 are essential for the consistent estimation of Γ and F_t. Specifically, these conditions ensure that the matrix $\Gamma' X_t' X_t \Gamma$, which is involved in inversion processes, remains nonsingular (where X_t denotes the $N \times L$ matrix consisting of the cross-section of $x_{i,t}$).

Assumption 4 *Asymptotic normality:*

(1) As $N, T \to \infty$, $\frac{1}{\sqrt{NT}} \sum_{i,t} vect\left(X_{i,t}' \epsilon_{i,t} F_t'\right) \xrightarrow{d} Normal\left(0, \Omega^{x\epsilon f}\right)$,

(2) As $N \to \infty$, $\frac{1}{\sqrt{N}} \sum_i vect\left(X_i' \epsilon_i\right) \xrightarrow{d} Normal\left(0, \Omega^{x\epsilon}\right)$ *for* $\forall t$,

(3) As $N, T \to \infty$, $\frac{1}{\sqrt{T}} \sum_t vect\left(F_t F_t' - E[F_t F_t']\right) \xrightarrow{d} Normal\left(0, \Omega^f\right)$.

Assumption 4 simply containes central limit theorems with respect to different variables, which are satisfied by various mixing processes.

3 Estimation

The GSC-IPCA estimator of the treatment effect for a treated unit i at time t is defined as the difference between the observed outcome and its estimated counterfactual: $\delta_{it} = Y_{it}(1) - \hat{Y}_{it}(0)$, where $\hat{Y}_{it}(0)$ is derived through a three-step imputation process.

Step 1: The initial step entails estimating the time-varying factors $\hat{F}_t$ and the coefficient matrix $\hat{\Gamma}_{\text{ctrl}}$ utilizing an Alternating Least Squares (ALS) algorithm, based exclusively on data from the control group.

$$(\hat{\Gamma}_{ctrl}, \hat{F}_t) = \arg\min_{\Gamma, F_t} \sum_{i \in N_{ctrl}} \sum_{t \in T} (Y_{it} - (X_{it}\Gamma)F_t)' (Y_{it} - (X_{it}\Gamma)F_t) . \tag{3}$$

Step 2: The subsequent step involves estimating the coefficient matrix $\hat{\Gamma}_{treat}$ for treated unit i at time t, employing the previously estimated time-varying factors $\hat{F}_t$ and the observed covariates X_{it}, using only pretreatment data from the treated units.

$$\hat{\Gamma}_{treat} = \arg\min_{\Gamma} \sum_{i \in N_{treat}} \sum_{t \in T_{pre}} \left(Y_{it} - (X_{it}\Gamma)\hat{F}_t\right)' \left(Y_{it} - (X_{it}\Gamma)\hat{F}_t\right) . \tag{4}$$

Step 3: The final step involves imputing the counterfactual outcome $\hat{Y}_{it}(0)$ for treated unit i at time t by substituting the estimated coefficient matrix $\hat{\Gamma}_{treat}$ and the time-varying factors $\hat{F}_t$ into the following equation:

$$\hat{Y}_{it}(0) = (X_{it}\hat{\Gamma}_{treat})\hat{F}_t, \quad \forall i \in N_{treat}, \forall t \in T_{post}. \tag{5}$$

4 Monte Carlo Simulation

In this section, we employ Monte Carlo simulations to assess the performance of the GSC-IPCA estimator in finite sample settings. We juxtapose the GSC-IPCA estimator against the GSC-IFE estimator, as introduced by Xu (2017), alongside other prominent methodologies in the realm of causal inference, such as the DID estimator and the original SC estimator. Our comparative analysis focuses on key metrics including bias, mean squared error (MSE), and coverage probability.

We initiate our analysis with a data generating process that incorporates $L = 10$ and $K = 3$ time-varying covariates and common factors, along with unit and time fixed effects:

$$Y_{it} = D_i\delta_t + X_{it}\Lambda + (X_{it}\Gamma)F_t + \alpha_i + \xi_t + \epsilon_{it}. \tag{6}$$

where $X_{it} = [x_{it}^1, \ldots, x_{it}^{10}]$ denotes the vector of time-varying covariates, which follows a VAR(1) process. $X_{it} = \mu_i + A_i X_{i,t-1} + \nu_{it}$, where A_i is a $L \times L$ variance-covariance matrix[3], The drift term μ_i equals 0 for control units and 2 for treated units,[4], and ν_{it} is a $L \times 1$ vector of i.i.d. standard normal errors. While $F_t = [f_t^1, \ldots, f_t^3]'$ denotes the vector of time-varying common factors, adhering to a similar VAR(1) process, the variable ϵ_{it} represents the idiosyncratic error term. Unit and time fixed effects, α_i and ξ_t respectively, are uniformly drawn from the interval $(0, 1)$. The coefficient vector $\Lambda = [\lambda^1, \ldots, \lambda^{10}]'$ associated with the covariates is drawn uniformly from $(0, 1)$, and Γ, the $L \times K$ coefficient matrix for the factor loadings, is drawn uniformly from $(0, 0.1)$, with these covariates serving as instruments. The treatment indicator D_{it} is binary, defined as $D_{it} = 1$ for treated units during post-treatment periods, and $D_{it} = 0$ otherwise. The heterogeneous treatment effect is modeled as

[3]In our methodology, the variance-covariance matrix is not constrained to be diagonal, thus allowing covariates within each unit to be correlated, reflecting the typical scenario in most economic time series data. To emphasize the independence among different units, we generate N unique variance-covariance matrices, each corresponding to a unit, ensuring cross-sectional independence and preserving time-series correlation. Moreover, we impose a condition on these matrices by requiring the eigenvalues of A_i to have characteristic roots that reside inside the unit circle, thereby assuring the stationarity of the VAR(1) process.

[4]This configuration underscores that the treatment assignment is not random; rather, it depends on the covariates X_{it}.

$\delta_{it} = \bar{\delta}_t + e_{it}$, where e_{it} is i.i.d as standard normal, and $\bar{\delta}_t = [0, \cdots, 0, 1, 2, \ldots, T_{post}]$ represents a time-varying treatment effect. Only the outcome Y_{it}, the covariates X_{it}, and the treatment indicator D_{it} are observed, while all other variables remain unobserved.

Table 1: Data Generating Process

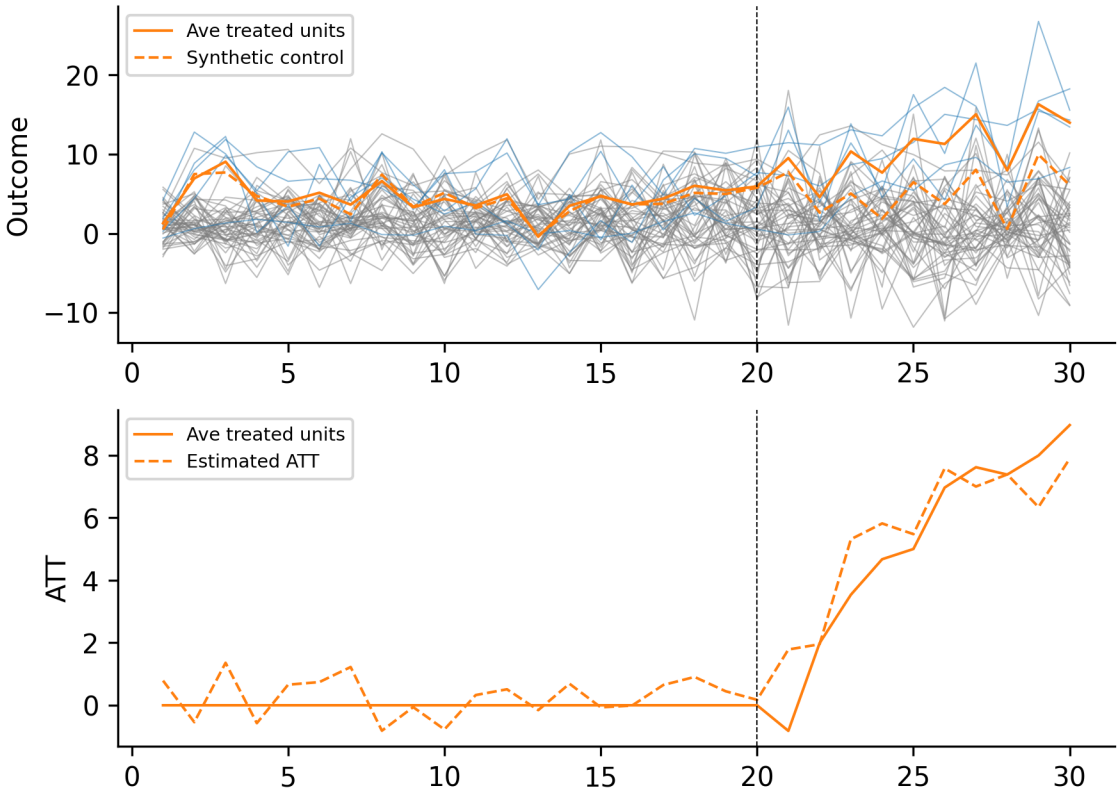
Variables	Description	Dimension	DGPs
D_i	treatment indicator	$N \times T$	$D_i = 1 \, \forall i \in N_{treat}$
δ_t	treatment effects	$N \times T$	$\bar{\delta}_t = [0, \cdots, 0, 1, 2, \ldots, T_{post}]$
X_{it}	covariates	$N \times T \times L$	$X_{it} = \mu_i + A_i X_{i,t-1} + \nu_{it}$
A_i	VAR(1) coefficients	$N \times L \times L$	$A_i \sim N(0,1)$, $\forall \lambda \in \text{eigen}(A_i)$, $\lvert\lambda\rvert \leq 1$
Λ	coefficients	$L \times 1$	$\Lambda \sim Uniform(0,1)$
Γ	coefficient matrix	$L \times K$	$\Gamma \sim Uniform(0,1)$
F_t	factors	$K \times 1$	$F_t = AF_{t-1} + \nu_t$
A	VAR(1) coefficients	$K \times K$	$A \sim N(0,1)$, $\forall \lambda \in \text{eigen}(A)$, $\lvert\lambda\rvert \leq 1$
α_i	unit FE	$N \times 1$	$\alpha_i \sim Uniform(0,1)$
ξ_t	time FE	$1 \times T$	$\xi_t \sim Uniform(0,1)$
ϵ_{it}	error term	$N \times T$	$\epsilon_{it} \sim N(0,1)$
Y_{it}	outcome	$N \times T$	Equation 6

4.1 A simulated example

In this simulated example, we demonstrate the efficacy of the GSC-IPCA estimator within a framework comprising $N_{\text{treat}} = 5$ treated units and $N_{\text{ctrl}} = 45$ control units, across $T_0 = 20$ pre-treatment and $T_1 = 10$ post-treatment periods. The simulation includes $L = 10$ covariates and $K = 3$ common factors. Figure 1 illustrates both the raw data and the imputed counterfactual outcomes as estimated by the GSC-IPCA method. In the upper panel, control units are represented in gray and treated units in light blue, with the average outcome for treated units highlighted in orange. The imputed synthetic average for treated outcomes is also shown, delineated by an orange dashed line. The GSC-IPCA method is capable of capturing the trajectory of the average outcome for treated units before treatement.

The lower panel of Figure 1 shows the estimated ATT (dashed line) with the true ATT (solid line). The GSC-IPCA method is able to capture the true ATT, as evidenced by the close alignment between the dashed and solid lines.

Figure 1: **GSC-IPCA estimation for simulated data**



This graphic plots the GSC-IPCA method estimated ATT for simulated data $N_{treat} = 5, N_{ctrl} = 45, T_0 = 20, T_1 = 10, L = 10$.

5 Empirical Application

In this section, we

6 Conclusion

Firms'

References

Yiqing Xu. Generalized synthetic control method: Causal inference with interactive fixed effects models. Political Analysis, 25(1):57–76, 2017.

A.1 Estimation of the GSC-IPCA Estimator Using the ALS Algorithm

As outlined in Equation 2, the data generating process can be described by:

$$Y_{it} = (X_{it}\Gamma)F_t + \epsilon_{it}, \quad \epsilon_{it} = \mu_{it} + H_{it}F_t.$$

Equation 3 details the derivation of the GSC-IPCA estimator from the minimization problem:

$$(\hat{\Gamma}, \hat{F}_t) = \arg\min_{\Gamma, F_t} \sum_{i \in N} \sum_{t \in T} (Y_{it} - (X_{it}\Gamma)F_t)' (Y_{it} - (X_{it}\Gamma)F_t).$$

The Alternating Least Squares (ALS) method is employed for the numerical solution of this optimization problem. Unlike PCA, the IPCA optimization challenge cannot be resolved through eigen-decomposition. The optimization, as defined in Equation 3, is quadratic with respect to either Γ or F_t, when the other is held constant. This characteristic permits the analytical optimization of Γ and F_t sequentially. With a fixed Γ, the solutions for F_t are t-separable and can be obtained via cross-sectional OLS for each t:

$$\hat{F}_t(\Gamma) = (\Gamma' X_t' X_t \Gamma)^{-1} \Gamma' X_t' Y_t.$$

Conversely, with known F_t, the optimal Γ (vectorized as γ) is derived through pooled panel OLS of y_{it} against LK regressors, $x_{it} \otimes f_t$:

$$\hat{\gamma} = \left(\sum_{i,t} (x_{i,t}' \otimes f_t)(x_{i,t} \otimes f_t') \right)^{-1} \left(\sum_{i,t} (x_{i,t}' \otimes f_t) y_{i,t} \right).$$

Inspired by PCA, the initial guess for F_t is the first K principal components of the outcome matrix Y_{it}. The ALS algorithm alternates between these two steps until convergence is achieved, typically reaching a local minimum rapidly. The convergence criterion, based on the relative change in the optimization problem from Equation 3, ensures termination when

this change falls below a predefined threshold, set at $10e^{-6}$ in our implementation.