Causal inference using factor models

Bai, Jushan and Wang, Peng

Columbia University, Hong Kong University of Science and Technology

31 March 2024

Causal Inference Using Factor Models

Jushan Bai*, Peng Wang† ‡

March, 2024

Abstract

We propose a framework for causal inference using factor models. We base our identification strategy on the assumption that policy interventions cause structural breaks in the factor loadings for the treated units. The method allows heterogeneous trends and is easy to implement. We compare our method with the synthetic control methods of Abadie, et al (2010, 2015), and obtain similar results. Additionally, we provide confidence intervals for the causal effects. Our approach expands the toolset for causal inference.

Key words: synthetic control, difference-in-differences, structural breaks, latent factors.

1 Introduction

Causal inference using synthetic control (SC) or difference-in-differences (DID) methods has become increasingly popular in the literature. The synthetic control method (Abadie and Gardeazabal (2003), Abadie, et al. (2010, 2015)) involves constructing a synthetic control unit that closely matches the treated unit in terms of pre-treatment characteristics and outcomes, and then comparing the outcomes of the treated unit with the synthetic control unit after the intervention. The difference-in-differences method (e.g., Card and Krueger (1994)) compares the change in outcomes for a treated group before and after the intervention with the change in outcomes for a control group over the same time period. Both methods rely on the assumption of a parallel trend, which means that the treated and control groups would have followed the same trend in the absence of the policy intervention. This assumption is used to construct

*Department of Economics, Columbia University, New York, NY, USA. Email: jb3064@columbia.edu.

†Department of Economics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: pwang@ust.hk.

‡We thank comments and suggestions from participants at the Conference on Time Series Econometrics and Model Checking at HKUST in 2023.

a counterfactual scenario where the policy intervention did not occur, and the difference between the observed outcomes and the counterfactual outcomes is defined as the causal effect of interest. However, in practice, individual heterogeneities in trends are common, which can violate the parallel trend assumption. For example, Gobillon and Magnac (2016) and Xu (2017) modeled such heterogeneous trends in panel data models with interactive fixed effects. The factor model acknowledges commonalities in comovement using common factors and uses unit-specific factor loadings to reflect individual heterogeneities. Using a factor model, we decompose the potential outcomes into a systematic component and an idiosyncratic component. The policy intervention affects the systematic component but not the idiosyncratic one. This allows us to represent the causal effect using the difference in the systematic components only. The advantage of using a factor model is that it allows us to clearly define and identify the causal effect. In addition, the role played by the policy intervention naturally translates into a structural-break problem, which can be tested using existing theory on structural breaks.

Next, we describe the model and demonstrate how this framework relates to existing approaches and how it facilitates the identification of causal effects.

2 Modeling the Potential Outcomes

We consider a panel data set that includes an outcome variable and some covariates. The observed outcome variable Y is indexed by unit and time, i.e., Y_{it}, $i = 1, 2, \ldots, n$, $t = 1, 2, \ldots, T$. The potential outcome for unit i in period t is denoted by $Y_{it}(d)$, $d = 0, 1$, with $d = 1$ referring to the case of treatment and $d = 0$ for the case of no treatment. Assume that a policy intervention occurs in period T_0 with $1 < T_0 < T$. Let D_{it} denote the observed treatment dummy. The policy intervention only applies to units $i \leq n_0$ without directly affecting units $i > n_0$. To focus on the main idea, we assume that the policy intervention occurs in the same period for all treated units. The treatment status can be summarized by

$$D_{it} = \begin{cases} 0, & i > n_0 \ \& \ 1 \leq t \leq T, \ \text{ or } \ i \leq n_0 \ \& \ t < T_0, \\ 1, & i \leq n_0 \ \& \ t \geq T_0. \end{cases} \tag{1}$$

The potential outcomes are assumed to follow a factor model:

$$Y_{it}(d) = \lambda_i(d)' f_t + X_{it}' \beta(d) + \varepsilon_{it}, \ d = 0, 1, \tag{2}$$

where X_{it} is the vector of observed covariates and ε_{it} denotes the idiosyncratic error that is not indexed by d. For the treated units we have,

$$Y_{it} = \begin{cases} \lambda_i(0)' f_t + X_{it}'\beta(0) + \varepsilon_{it} = Y_{it}(0), & t < T_0, \\ \lambda_i(1)' f_t + X_{it}'\beta(1) + \varepsilon_{it} = Y_{it}(1), & t \geq T_0, \end{cases}, \; i = 1, \ldots, n_0.$$

For the untreated units

$$Y_{it} = \lambda_i(0)' f_t + X_{it}'\beta(0) + \varepsilon_{it} = Y_{it}(0), \; i = n_0 + 1, \ldots, n, \; t = 1, ..., T.$$

The causal effect for the treated unit is

$$\tau_{it} = Y_{it}(1) - Y_{it}(0) = Y_{it} - Y_{it}(0), \; t \geq T_0, \; i = 1, \ldots, n_0.$$

Using the factor model, we can rewrite τ_{it} as ($i \leq n_0$, $t > T_0$)

$$\begin{aligned} \tau_{it} &= \{\lambda_i(1)' f_t + X_{it}'\beta(1) + \varepsilon_{it}\} - \{\lambda_i(0)' f_t + X_{it}'\beta(0) + \varepsilon_{it}\} \\ &= [\lambda_i(1) - \lambda_i(0)]' f_t + X_{it}' [\beta(1) - \beta(0)]. \end{aligned} \tag{3}$$

This representation allows us to evaluate the source of the causal effects due to structural breaks in factor loadings, or covariates' coefficients, or both.

Table 1 provides a timeline of the potential and observed outcomes. For the control group ($j > n_0$), the observed outcome Y_{jt} equals the potential outcome $Y_{jt}(0)$ for all t. For the treatment group ($i \leq n_0$), the observed outcome Y_{it} equals the potential outcome $Y_{it}(0)$ before the intervention ($t < T_0$) and equals the potential outcome $Y_{it}(1)$ post intervention ($t \geq T_0$). The last two rows of Table 1 give, respectively, the counterfactual and treatment effects for the treated group.

Figure 1 provides an example of the relationship among factors, realized individual trend before the intervention ($\lambda_i(0)' \cdot f_s$, $s < T_0$) and after the intervention ($\lambda_i(1)' \cdot f_t$, $t \geq T_0$), as well as the potential individual trend ($\lambda_i(0)' \cdot f_t$, $t \geq T_0$). In the figure, f_t is illustrated as a smooth function of t, but it does not have to be smooth. Also, only a single factor is illustrated. The symbols `L(0)` and `L(1)` represents $\lambda_i(0)$ and $\lambda_i(1)$, respectively. To focus on the main idea, we abstract from the covariates in Table 1 and Figure 1.

In Section 3, we show that the individual causal effect τ_{it} in (3) is identifiable because $\{\lambda_i(d), f_t, \beta(d)\}$, $d = 0, 1$, are all identifiable. Then a natural estimator for τ_{it} is given by

$$\hat{\tau}_{it} = \left[\hat{\lambda}_i(1) - \hat{\lambda}_i(0)\right]' \hat{f}_t + X_{it}' \left[\hat{\beta}(1) - \hat{\beta}(0)\right], \tag{4}$$

where the hatted variables are the corresponding estimates.

	$s < T_0$	$t \geq T_0$
Treated $(i \leq n_0)$	$Y_{is} = \underbrace{\lambda_i (0)' f_s + \varepsilon_{is}}_{Y_{is}(0)}$	$Y_{it} = \underbrace{\lambda_i (1)' f_t + \varepsilon_{it}}_{Y_{it}(1)}$
Control $(j > n_0)$	$Y_{js} = \underbrace{\lambda_j (0)' f_s + \varepsilon_{js}}_{Y_{js}(0)}$	$Y_{jt} = \underbrace{\lambda_j (0)' f_t + \varepsilon_{jt}}_{Y_{jt}(0)}$
Counterfactual $(i \leq n_0)$		$Y_{it} (0) = \lambda_i (0)' f_t + \varepsilon_{it}$
Causal effect $(i \leq n_0)$		$\tau_{it} = \underbrace{\lambda_i (1)' f_t - \lambda_i (0)' f_t}_{Y_{it}(1) - Y_{it}(0)}$

Table 1: Outcomes before and after the intervention.

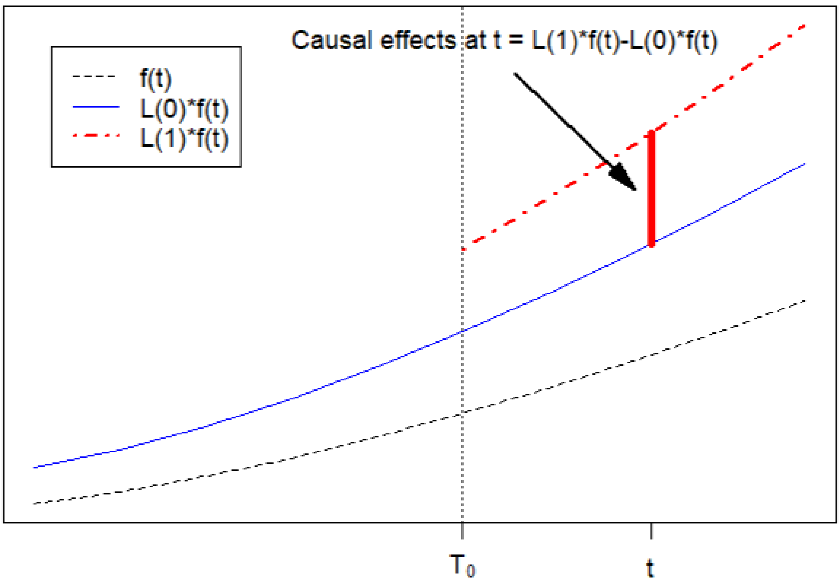

Figure 1: Individual trend before and after the intervention.

We will show that

$$\hat{\tau}_{it} - \tau_{it} = o_p(1),$$

under standard conditions such as the ones in Bai (2009). Such a model-based causal model allows us to identify the causal effect using the systematic component regardless of the idiosyncratic errors. Existing causal inference methods, such as difference-in-differences, synthetic control, and matrix completion (e.g., Bai and Ng (2021)), focus on constructing the counterfactual $Y_{it}(0)$ from the control group. Our model-based method does not directly construct $Y_{it}(0)$, noting that the idiosyncratic errors will cancel out in the difference $Y_{it}(1) - Y_{it}(0)$.

2.1 Compare with the Causal Model using Interactive Fixed Effects

The causal model that we propose is closely related to Gobillon and Magnac (2016) and Xu (2017). Both of their models can be summarized by the following model for potential outcomes

$$\begin{aligned} Y_{it}(0) &= \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}, \\ Y_{it}(1) &= \delta_{it} + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}, \end{aligned}$$

where $\delta_{it} = Y_{it}(1) - Y_{it}(0)$ is defined as the individual causal effect. The estimator for this causal effect is given by

$$\begin{aligned} \hat{\delta}_{it} &= Y_{it}(1) - \hat{Y}_{it}(0) \\ &= Y_{it} - \left(\hat{\lambda}_i' \hat{f}_t + X_{it}'\hat{\beta}\right), \quad t \geq T_0,\ i = 1, \ldots, n_0. \end{aligned} \tag{5}$$

Plug-in the model for $Y_{it}(1)$ to obtain

$$\begin{aligned} \hat{\delta}_{it} &= \left(\delta_{it} + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}\right) - \left(\hat{\lambda}_i' \hat{f}_t + X_{it}'\hat{\beta}\right) \\ &= \delta_{it} + \left(\lambda_i' f_t - \hat{\lambda}_i' \hat{f}_t\right) + X_{it}'\left(\beta - \hat{\beta}\right) + \varepsilon_{it}. \end{aligned}$$

Under appropriate assumptions such as the ones in Bai (2009), $\lambda_i' f_t - \hat{\lambda}_i' \hat{f}_t = o_p(1)$ and $\beta - \hat{\beta} = o_p(1)$. So

$$\hat{\delta}_{it} - \delta_{it} = o_p(1) + \varepsilon_{it} = O_p(1). \tag{6}$$

As a result, under Gobillon and Magnac (2016) and Xu (2017)'s modeling strategy, the idiosyncratic error ε_{it} strongly affects the bias in the individual causal estimates. The bias could be averaged out when the number of treated units, n_0, is large. However, the framework could be problematic when n_0 is small, as in most applications in the synthetic control literature. Our modeling strategy directly acknowledges the possible variations of factor loadings and covariate coefficients across different treatment status. Our estimator of the individual causal effect is given by (4)

$$\hat{\tau}_{it} = \left[\hat{\lambda}_i(1) - \hat{\lambda}_i(0)\right]' \hat{f}_t + X_{it}' \left[\hat{\beta}(1) - \hat{\beta}(0)\right],$$

which direct corresponds to the model-implied causal effect τ_{it} in (3) . The idiosyncratic error ε_{it} itself does not lead to a bias in $\hat{\tau}_{it} - \tau_{it}$. In addition, our modeling strategy allows n_0 to be either large or small, and thus is more suitable for studying the cases when one or only a few units are subject to the policy intervention.

Similarly, Callaway and Karami (2023) propose to model the untreated potential outcomes as

$$Y_{it}(0) = \theta_t + \xi_i + \lambda_i' f_t + \varepsilon_{it}.$$

which focuses on the average treatment effect. Their framework is helpful for the applications in the difference-in-differences literature but is not designed for the small n_0 case as in the synthetic control literature.

More importantly, Gobillon and Magnac (2016), Xu (2017), and Callaway and Karami (2023) all focus on modeling the untreated potential outcomes $Y_{it}(0)$ and impose no restrictions on the causal effect δ_{it}. They do not have an explicit model for $Y_{it}(1)$. The cost of such a flexibility is the bias in the estimator of individual causal effects as shown in (6). In comparison, we explicitly model both $Y_{it}(0)$ and $Y_{it}(1)$, which constraints the causal effect $\delta_{it} = \tau_{it}$ through equation (3). The implied estimator (4) for individual causal effect is free of bias when sample sizes are large, even when the number of treated units is small.

2.2 Compare with the Difference-in-Differences Model using Interactive Fixed Effects

Here we show that the difference-in-differences method with or without interactive fixed effects is a special case of our set up. A standard difference-in-differences model with constant treatment effects and a common treatment timing features the following two-way-fixed-effects (TWFE) panel regression model:

$$Y_{it} = \alpha_i + \theta_t + \rho D_{it} + X_{it}'\beta + \varepsilon_{it}, \tag{7}$$

with D_{it} being the treatment indicator defined in (1). Adding the interactive fixed effects, (7) can be generalized to:

$$Y_{it} = \alpha_i + \theta_t + \rho D_{it} + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}. \tag{8}$$

In both (7) and (8), ρ represents the treatment effect.

The corresponding potential outcomes are given by

$$Y_{it}(0) = \alpha_i + \theta_t + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it},$$
$$Y_{it}(1) = \alpha_i + \theta_t + \rho + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}. \tag{9}$$

Define $g_t = [1, \theta_t, f_t']'$, $\lambda_i(0) = [\alpha_i, 1, \lambda_i']'$, $\lambda_i(1) = [\alpha_i + \rho, 1, \lambda_i']'$, and then we may represent (9) as

$$Y_{it}(0) = \lambda_i(0)' g_t + X_{it}'\beta + \varepsilon_{it},$$
$$Y_{it}(1) = \lambda_i(1)' g_t + X_{it}'\beta + \varepsilon_{it}, \tag{10}$$

which is a special case of our causal model (2) in which $\beta(d) = \beta,\ d = 0, 1$. Our representation of the causal effect (3) takes into account variations in λ_i and β across different treatment status to model the heterogeneous causal effect.

If one wants to use TWFE model to learn about the heterogeneous causal effects, the regression model can be specified as

$$Y_{it} = \alpha_i + \theta_t + \rho_i D_{it} + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}. \tag{11}$$

where the coefficient of D_{it} is individual-dependent. The corresponding potential outcomes are given by

$$Y_{it}(0) = \alpha_i + \theta_t + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it},$$
$$Y_{it}(1) = \alpha_i + \theta_t + \rho_i + \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}. \tag{12}$$

Define $g_t = [1, \theta_t, f_t']'$, $\lambda_i(0) = [\alpha_i, 1, \lambda_i']'$, $\lambda_i(1) = [\alpha_i + \rho_i, 1, \lambda_i']'$, and then we may represent (12) as (10), again a special case of our causal model (2).

2.3 Compare with the Synthetic Control Method

To deliver the main idea, assume that the potential outcomes follow a factor model without covariates

$$Y_{it}(d) = \lambda_i(d)' f_t + \varepsilon_{it},\ \ d = 0, 1, \tag{13}$$

Assume $i = 1$ is treated with a policy intervention in period T_0, and the unaffected control units are $i = 2, ..., n$. The synthetic control method constructs the counterfactual $Y_{1t}(0)$ as a weighted average of the observed outcomes for the control units:

$$\widehat{Y}_{1t}(0) = \sum_{i=2}^{n} \omega_i Y_{it}, \quad t \geq T_0,\ \omega_i \geq 0,\ \sum_{i=2}^{n} \omega_i = 1.$$

Then

$$\begin{aligned}\widehat{Y}_{1t}(0) &= \sum_{i=2}^{n} \omega_i \left(\lambda_i(0)' f_t + \varepsilon_{it}\right) \\ &= \left(\sum_{i=2}^{n} \omega_i \lambda_i(0)'\right) f_t + \sum_{i=2}^{n} \omega_i \varepsilon_{it}.\end{aligned}$$

The synthetic causal effect for $t \geq T_0$ is

$$\begin{aligned}\tau_{1t}^{synth} &\equiv Y_{1t} - \widehat{Y}_{1t}(0) \\ &= \lambda_1(1)' f_t + \varepsilon_{1t} - \left\{\left(\sum_{i=2}^{n} \omega_i \lambda_i(0)'\right) f_t + \sum_{i=2}^{n} \omega_i \varepsilon_{it}\right\} \\ &= \left[\lambda_1(1) - \sum_{i=2}^{n} \omega_i \lambda_i(0)\right]' f_t + \left[\varepsilon_{1t} - \sum_{i=2}^{n} \omega_i \varepsilon_{it}\right].\end{aligned}$$

Our model-based causal effect is given by

$$\tau_{1t}^{factor} \equiv [\lambda_1(1) - \lambda_1(0)]' f_t, \quad t \geq T_0.$$

The difference between this two is

$$\tau_{1t}^{synth} - \tau_{1t}^{factor} = \left[\lambda_1(0) - \sum_{i=2}^{n} \omega_i \lambda_i(0)\right]' f_t + \left[\varepsilon_{1t} - \sum_{i=2}^{n} \omega_i \varepsilon_{it}\right],\ t \geq T_0.$$

The synthetic control chooses the weights such that the distance between Y_{1t} and $\sum_{i=2}^{n} \omega_i Y_{it}$ is small for $t < T_0$. Note that for $t < T_0$,

$$\begin{aligned}Y_{1t} - \sum_{i=2}^{n} \omega_i Y_{it} &= Y_{1t}(0) - \sum_{i=2}^{n} \omega_i Y_{it}(0) \\ &= \lambda_1(0)' f_t + \varepsilon_{1t} - \sum_{i=2}^{n} \omega_i \left[\lambda_i(0)' f_t + \varepsilon_{it}\right] \\ &= \left[\lambda_1(0) - \sum_{i=2}^{n} \omega_i \lambda_i(0)\right]' f_t + \left[\varepsilon_{1t} - \sum_{i=2}^{n} \omega_i \varepsilon_{it}\right].\end{aligned}$$

In the ideal case that $\omega_i' s$ are chosen such that

$$\lambda_1(0) \approx \sum_{i=2}^{n} \omega_i \lambda_i(0),$$

we have

$$\tau_{1t}^{synth} - \tau_{1t}^{factor} \approx \varepsilon_{1t} - \sum_{i=2}^{n} \omega_i \varepsilon_{it},\ t \geq T_0.$$

If we expect that the factors already reflect most of the cross-sectional correlations, the correlation between ε_{1t} and $\sum_{i=2}^{n} \omega_i \varepsilon_{it}$ is weak. In the special case where $\{\varepsilon_{it}\}$ is iid. across i, $\varepsilon_{1t} - \sum_{i=2}^{n} \omega_i \varepsilon_{it} = \varepsilon_{1t} + o_p(1) = O_p(1)$ and thus $\tau_{1t}^{synth} - \tau_{1t}^{factor} = O_p(1)$. The difference can be small when averaged over time. The difference will also be small when averaged over many treated units. In general, however, the difference may not be negligible.

Hsiao, et al (2012) adopted a similar method as synthetic control. They start with a factor model and propose to use a linear function of outcomes for the untreated units to estimate the counterfactual $Y_{1t}(0)$, $t \geq T_0$. Accordingly, the estimator for the causal effect is

$$Y_{1t} - \hat{Y}_{1t}(0), \ t \geq T_0.$$

A simple regression method is used to estimate the optimal linear function. Both our method and Hsiao, et al (2012)'s do not require numerical optimization and are easy to implement.

3 Identification of the Causal Factor Model

3.1 The Benchmark Model without Covariates

The causal effect for the treated unit is defined as

$$\tau_{it} = Y_{it}(1) - Y_{it}(0) = Y_{it} - Y_{it}(0), \ \ i \leq n_0, \ t \geq T_0.$$

To focus on the main idea, we first study the case without covariates. Plug in the factor model to obtain

$$\begin{aligned} \tau_{it} &= \{\lambda_i(1)' f_t + \varepsilon_{it}\} - \{\lambda_i(0)' f_t + \varepsilon_{it}\} \\ &= [\lambda_i(1) - \lambda_i(0)]' f_t, \ \ i \leq n_0, \ t \geq T_0. \end{aligned}$$

This individual causal effect τ_{it} is identifiable as $\{\lambda_i(1), \lambda_i(0), f_t\}$ are all identified.

- The factors f_t $(1 \leq t \leq T)$ can be identified using existing methods such as the principal component analysis of the untreated units $\{Y_{it}\}$, $i > n_0$, $t = 1, \ldots, T$.
- $\lambda_i(0)$ is identified by regressing Y_{it} on f_t for $t < T_0$, $i \leq n_0$.
- $\lambda_i(1)$ is identified by regressing Y_{it} on f_t for $t \geq T_0$, $i \leq n_0$.
- The test for $H_0 : \tau_{it} = 0, \ t \geq T_0, \ i \leq n_0$ is the same as testing $H_0 : \lambda_i(1) = \lambda_i(0), \ i \leq n_0$. This can be the structural break test for the equation $Y_{it} = \lambda_i' f_t + \varepsilon_{it}, \ i \leq n_0, \ 1 \leq t \leq T$.

Note that the identification strategy works for the cases when n_0 is either small or large. In particular, it works for the special case where $n_0 = 1$, similar to the synthetic control setup.

3.2 The Model with Covariates (1)

Assuming that the coefficients of covariates do not depend on treatment status, the data generating process is

$$Y_{it}(d) = \lambda_i(d)' f_t + X_{it}'\beta + \varepsilon_{it}.$$

Then we have,

$$\begin{aligned} Y_{it} &= Y_{it}(0), \; i > n_0, \; 1 \leq t \leq T, \\ Y_{it} &= \begin{cases} Y_{it}(0), & i \leq n_0, \; t < T_0, \\ Y_{it}(1), & i \leq n_0, \; t \geq T_0. \end{cases} \end{aligned}$$

The causal effect for the treated unit is

$$\tau_{it} = Y_{it}(1) - Y_{it}(0) = Y_{it} - Y_{it}(0), \;\; i \leq n_0, \; t \geq T_0.$$

Plug in the factor model, we have

$$\begin{aligned} \tau_{it} &= \left\{\lambda_i(1)' f_t + X_{it}'\beta + \varepsilon_{it}\right\} - \left\{\lambda_i(0)' f_t + X_{it}'\beta + \varepsilon_{it}\right\} \\ &= \left[\lambda_i(1) - \lambda_i(0)\right]' f_t, \;\; i \leq n_0, \; t \geq T_0. \end{aligned}$$

This individual causal effect τ_{it} is identifiable as $\{\lambda_i(1), \lambda_i(0), f_t\}$ are all identified.

- The factors f_t $(1 \leq t \leq T)$ can be identified using existing methods such as the panel regression with interactive fixed effects of the untreated units $\{Y_{it}\}$, $i > n_0, \; t = 1, \ldots, T$.
  - As a by-product, β is also identified.
- $\lambda_i(0)$ is identified by regressing Y_{it} on f_t and X_{it} for $i \leq n_0, t < T_0$.
- $\lambda_i(1)$ is identified by regressing Y_{it} on f_t and X_{it} for $i \leq n_0, t \geq T_0$.
- The test for $H_0 : \tau_{it} = 0, \; t \geq T_0, \; i \leq n_0$ is the same as testing $H_0 : \lambda_i(1) = \lambda_i(0), \; i \leq n_0$. This is the structural break test for the equation $Y_{it} = \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it}, \; i \leq n_0, \; 1 \leq t \leq T$.

3.3 The Model with Covariates (2)

When the policy intervention affects how the covariates influence potential outcomes, the data generating process becomes

$$Y_{it}(d) = \lambda_i(d)' f_t + X_{it}'\beta(d) + \varepsilon_{it},$$

where the coefficient β depends on the treatment status. The causal effect for the treated unit is still defined as

$$\tau_{it} = Y_{it}(1) - Y_{it}(0) = Y_{it} - Y_{it}(0), \quad i \le n_0,\ t \ge T_0.$$

Plug in the factor model, we have

$$\begin{aligned}\tau_{it} &= \left\{\lambda_i(1)' f_t + X_{it}'\beta(1) + \varepsilon_{it}\right\} - \left\{\lambda_i(0)' f_t + X_{it}'\beta(0) + \varepsilon_{it}\right\} \\ &= \left[\lambda_i(1) - \lambda_i(0)\right]' f_t + X_{it}'\left[\beta(1) - \beta(0)\right], \quad i \le n_0,\ t \ge T_0.\end{aligned}$$

This individual causal effect τ_{it} is identifiable as $\{\lambda_i(1), \lambda_i(0), f_t, \beta(1), \beta(0)\}$ are all identified.

- The factors f_t ($1 \le t \le T$) can be identified using existing methods such as the panel regression with interactive fixed effects of the untreated units $\{Y_{it}\}$, $i > n_0,\ t = 1, \ldots, T$.
  - As a by-product, $\beta(0)$ is also identified.
- $\lambda_i(0)$ is identified by regressing Y_{it} on f_t and X_{it} for $i \le n_0, t < T_0$.
- $\lambda_i(1)$ and $\beta(1)$ is identified by regressing Y_{it} on f_t and X_{it} for $i \le n_0, t \ge T_0$.
- The test for $H_0 : \tau_{it} = 0,\ t \ge T_0,\ i \le n_0$ is the same as testing $H_0 : \lambda_i(1) = \lambda_i(0), i \le n_0$, and $\beta(1) = \beta(0)$. This is the structural break test for the equation $Y_{it} = \lambda_i' f_t + X_{it}'\beta + \varepsilon_{it},\ i \le n_0,\ 1 \le t \le T$.

4 Extension to the Case with Potential Factors

Assume $n_0/n \to c \in (0,1)$. In this case, we can identify a more flexible model where the factors for the treated group are affected by the intervention. Assume that the same set of common factors affect all units before the intervention. However, the policy intervention can affect the common factors for the treated group. For the untreated group, the common factors are given by $f_t(0), t = 1, 2, ..., T$, and for the treated group, the common factors are given by $f_t(0)$ for $t < T_0$, and by $f_t(1)$ for $t \ge T_0$.

For the treated group post intervention ($t \geq T_0$), we may think of $f_t(0)$ as its potential factors if the intervention never occurred, while $f_t(1)$ as the realized factors given that the intervention occurred at T_0. The data generating process for the potential outcome is

$$Y_{it}(d) = \lambda_i(d)' f_t(d) + \varepsilon_{it}, \quad d = 0, 1.$$

The observed outcome is given by,

$$\begin{aligned} Y_{it} &= Y_{it}(0), \ i > n_0, \ 1 \leq t \leq T, \\ Y_{it} &= \begin{cases} Y_{it}(0), & t < T_0, \\ Y_{it}(1), & t \geq T_0, \end{cases}, \quad i \leq n_0. \end{aligned}$$

The causal effect for the treated unit is

$$\tau_{it} = Y_{it}(1) - Y_{it}(0) = Y_{it} - Y_{it}(0), \ \ t \geq T_0, \ i \leq n_0.$$

Plug in the factor model, we have

$$\begin{aligned} \tau_{it} &= \left\{ \lambda_i(1)' f_t(1) + \varepsilon_{it} \right\} - \left\{ \lambda_i(0)' f_t(0) + \varepsilon_{it} \right\} \\ &= \lambda_i(1)' f_t(1) - \lambda_i(0)' f_t(0), \ \ t \geq T_0, \ i \leq n_0. \end{aligned} \tag{14}$$

This representation allows us to evaluate the source of the causal effects due to structural breaks in both factor loadings and factors.

The individual causal effect τ_{it} is identifiable as $\{\lambda_i(1), f_t(1), \lambda_i(0), f_t(0)\}$, $i \leq n_0, \ t \geq T_0$, are all identified. A simple identification strategy is given as follows.

- The factors $f_t(0)$ ($1 \leq t \leq T$) can be identified using principal component analysis of the untreated units$\{Y_{it}\}$, $i > n_0, \ t = 1, \ldots, T$.
- $\lambda_i(0)$, $i \leq n_0$, is identified by regressing Y_{it} on $f_t(0)$ for $t < T_0$.
- The counter factual $\lambda_i(0)' f_t(0)$, $i \leq n_0, \ t \geq T_0$, can be thus constructed as the product of the above two.
- The product $\lambda_i(1)' f_t(1)$ is identified by the principal component analysis of Y_{it}, $i \leq n_0, \ t \geq T_0$.

Let the estimators be $\left\{ \hat{\lambda}_i(1)' \hat{f}_t(1), \hat{\lambda}_i(0), \hat{f}_t(0) \right\}$, $i \leq n_0, \ t \geq T_0$. Then the estimator of the causal effect is given by

$$\hat{\tau}_{it} = \hat{\lambda}_i(1)' \hat{f}_t(1) - \hat{\lambda}_i(0)' \hat{f}_t(0), \ \ t \geq T_0, \ i \leq n_0. \tag{15}$$

Under standard identification constraints for factor models, such as Bai and Ng (2013),

$$\hat{\lambda}_i(d) \xrightarrow{p} \lambda_i(d), \; \hat{f}_t(d) \xrightarrow{p} f_t(d), \; t \geq T_0, \; i \leq n_0, \; d = 0, 1.$$

And thus we have

$$\hat{\tau}_{it} - \tau_{it} = o_p(1).$$

The asymptotic variance of $\hat{\tau}_{it}$ is considered below.

5 Estimation and Inference

5.1 The Intervention does not Affect the Factors

Consider (4) as the estimator for the unit-specific causal effect (3):

$$\hat{\tau}_{it} = \left[\hat{\lambda}_i(1) - \hat{\lambda}_i(0)\right]' \hat{f}_t + X'_{it}\left[\hat{\beta}(1) - \hat{\beta}(0)\right], \; t \geq T_0, \; i \leq n_0.$$

The factor estimate $\hat{f}_t$ is obtained using factor analysis of the control units. Then $\hat{\lambda}_i(0)$ and $\hat{\beta}(0)$ are obtained from a regression of y_{it} on $\hat{f}_t$ and X_{it} for $t < T_0$ and $i \leq n_0$. $\hat{\lambda}_i(1)$ and $\hat{\beta}(1)$ are obtained from another regression of y_{it} on $\hat{f}_t$ and X_{it} for $t \geq T_0$ and $i \leq n_0$. The large sample theory is a natural extension of Bai and Ng (2006), which is summarized in Proposition 1.

Proposition 1. Under Assumptions A1-A5 in the Appendix and $\sqrt{n - n_0}/T \to 0$, $\hat{\tau}_{it}$ is a consistent estimator of τ_{it} for $i \leq n_0$ and $t \geq T_0$, and

$$\frac{\hat{\tau}_{it} - \tau_{it}}{\sqrt{var(\hat{\tau}_{it})}} \xrightarrow{d} N(0, 1),$$

where

$$var(\hat{\tau}_{it}) = \hat{z}'_{it} \cdot var(\hat{\delta}_i)\, \hat{z}_{it} + \hat{\alpha}'_i \cdot var\left(\hat{f}_t\right) \hat{\alpha}_i,$$

and $\hat{z}_{it} = \left[\hat{f}'_t, X'_{it}\right]'$, $\hat{\alpha}_i = \hat{\lambda}_i(1) - \hat{\lambda}_i(0)$, $\hat{\delta}_i = \left[\hat{\alpha}'_i, \left[\hat{\beta}(1) - \hat{\beta}(0)\right]'\right]'$.

Proof. See Appendix. □

From Bai (2003), an estimator for the variance of the factor is given by

$$\widehat{var\left(\hat{f}_t\right)} = \frac{1}{n}\hat{V}^{-1}\hat{\Gamma}_t\hat{V}^{-1},$$

where $\hat{V}$ is the $r \times r$ diagonal matrix consisting of the r largest eigenvalues of $\frac{YY'}{(n-n_0)T}$

with Y being the observed $T \times (n - n_0)$ outcome matrix for the control units, and

$$\hat{\Gamma}_t = \frac{1}{n - n_0} \sum_{i=n_0+1}^{n} \hat{\varepsilon}_{it}^2 \hat{\lambda}_i \hat{\lambda}_i'.$$

We may define $\delta_i(d) = \left[\lambda_i(d)', \beta(d)'\right]'$, $d = 0, 1$. Then $\hat{\delta}_i = \hat{\delta}_i(1) - \hat{\delta}_i(0)$. In the Appendix, we show that

$$var\left(\hat{\delta}_i\right) = var\left(\hat{\delta}_i(1)\right) + var\left(\hat{\delta}_i(0)\right) + o_p(1),$$

and an estimator for $var\left(\hat{\delta}_i(d)\right)$, $d = 0, 1$, is given by

$$\frac{1}{T(d)}\left(\frac{1}{T(d)}\hat{z}_i(d)'\hat{z}_i(d)\right)^{-1}\left(\frac{1}{T(d)}\sum_{t=1}^{T(d)}\hat{\varepsilon}_{it}(d)^2\hat{z}_{it}(d)\hat{z}_{it}(d)'\right)\left(\frac{1}{T(d)}\hat{z}_i(d)'\hat{z}_i(d)\right)^{-1}.$$

5.2 The Intervention Affects the Factors for the Treated Group

Consider (15) as the estimator for the unit-specific causal effect (14):

$$\hat{\tau}_{it} = \hat{\lambda}_i(1)'\hat{f}_t(1) - \hat{\lambda}_i(0)'\hat{f}_t(0), \ t \geq T_0, \ i \leq n_0.$$

The factor estimate $\hat{f}_t(0)$ is obtained using factor analysis of the control units. Then $\hat{\lambda}_i(0)$ is obtained from a regression of y_{it} on $\hat{f}_t$ for $t < T_0$ and $i \leq n_0$. The product $\hat{\lambda}_i(1)'\hat{f}_t(1)$ is obtained as the common component estimator from principal component analysis of y_{it} for $t \geq T_0$ and $i \leq n_0$. We have the following proposition.

Proposition 2. Under Assumptions A1-A5 in the Appendix and $n_0/n \to c \in (0,1)$, $T_0/T \to b \in (0,1)$, $\sqrt{n}/T \to 0$, $\hat{\tau}_{it}$ is a consistent estimator of τ_{it} for $i \leq n_0$ and $t \geq T_0$ and

$$V_{it}^{-1/2}\left(\hat{\tau}_{it} - \tau_{it}\right) \xrightarrow{d} N(0,1),$$

for some $V_{it} > 0$.

Proof. See Appendix. □

We will also provide an estimator for V_{it} in the Appendix.

6 Two Empirical Applications

6.1 An Application to Abadie, et al (2010)

Using Abadie, et al (2010)'s data on per capita cigarette sales among 39 US states, we construct the counterfactual California using factor models and compare it with the

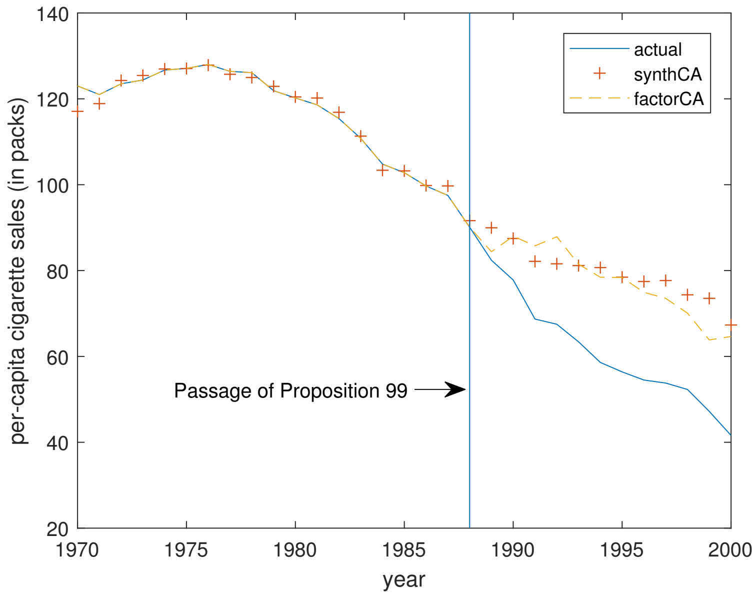

Figure 2: Causal factor model vs. synthetic control: the counterfactual California

synthetic-California. We find that the two approaches yield very close results. Bai and Ng (2002)'s criteria found two common factors among the 38 control states. We also tried to estimate one or three factors and obtain similar results. Such robustness comes from the fact that the first factor explains the majority of the variations in the data. We proceed in the following steps.

- Step 1: use principal component analysis for 38 control states to obtain the factor estimates $\hat{f}_t$.
- Step 2: regress Y_{it} on $\hat{f}_t$ for $t < T_0$ to obtain $\hat{\lambda}_i(0)$, $i = CA$. Regress Y_{it} on $\hat{f}_t$ for $t \geq T_0$ to obtain $\hat{\lambda}_i(1)$ and $\hat{\varepsilon}_{it}$, $i = CA$.
- Step 3: the estimator for the causal effect is

$$\hat{\tau}_{CA,t} = \left[\hat{\lambda}_{CA}(1) - \hat{\lambda}_{CA}(0)\right]\hat{f}_t,\ t \geq T_0.$$

In Figure 2, we compare the actual data, the counterfactual CA using either factor models or synthetic control method. The vertical line represents year 1988, which was the year of the passage of Proposition 99 in California. The treatment periods are from $T_0 = 1989$ to $T = 2000$. The counterfactual CA using factor models is defined as $\hat{Y}_{CA,t}(0) = \hat{\lambda}_{CA}(0)\hat{f}_t + \hat{\varepsilon}_{CA,t}$, $t \geq T_0$, where $\hat{\varepsilon}_{CA,t}$ is obtained in Step 2. Figure 3 shows the estimated causal effects using either factor models or synthetic control. To summarize, these two approaches yield very similar results.

We further investigate whether the policy intervention indeed induced a structural break by regressing California's observed outcome on the two factors using the whole

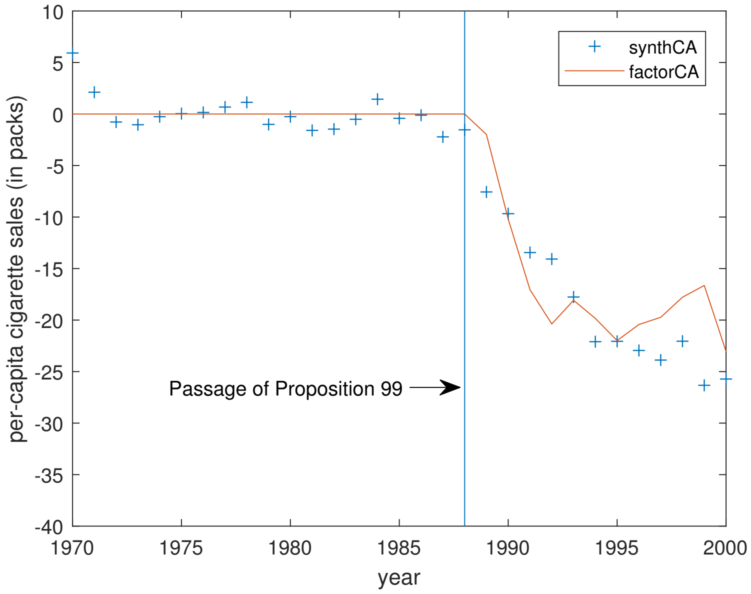

Figure 3: Causal factor model vs. synthetic control: the causal effects

sample. The Quandt Likelihood Ratio test for a structural break at an unknown point with 15 percent trimming yields a p-value of 0.0000, with the maximum F-statistic at observation 1993. In the meantime, the Chow-test for a structural break at observation 1989 yields an F-statistic of 21.26 with p-value 0.0000. All such evidence supports our proposal of using time-varying factor loadings to model the causal effect from an intervention.

Applying the results from Section 5, we may construct the 95% confidence interval of our causal estimates based on the factor models. Figure 4 reproduces Figure 3 with the shaded region being the 95% confidence interval around the causal estimates. The confidence intervals show that our causal estimates are mostly significant at 5% level. In addition, the confidence interval covers the causal estimates from the synthetic control method.

6.2 An Application to Abadie, et al (2015)

In this section, we compare the causal estimates from the synthetic control method and causal factor models using Abadie, et al (2015)'s data on per capita GDP for 17 countries. The objective is to evaluate the causal impact of German re-unification on Germany's per capita GDP. The synthetic control method uses 16 countries to construct the synthetic West Germany. The causal factor model uses the same 16 countries to construct the counterfactual West Germany. We maintain the assumption that German re-unification did not have impact on the other 16 countries.

In Figure 5, time series of per capita GDP for 17 countries (1969 - 2003) demonstrate

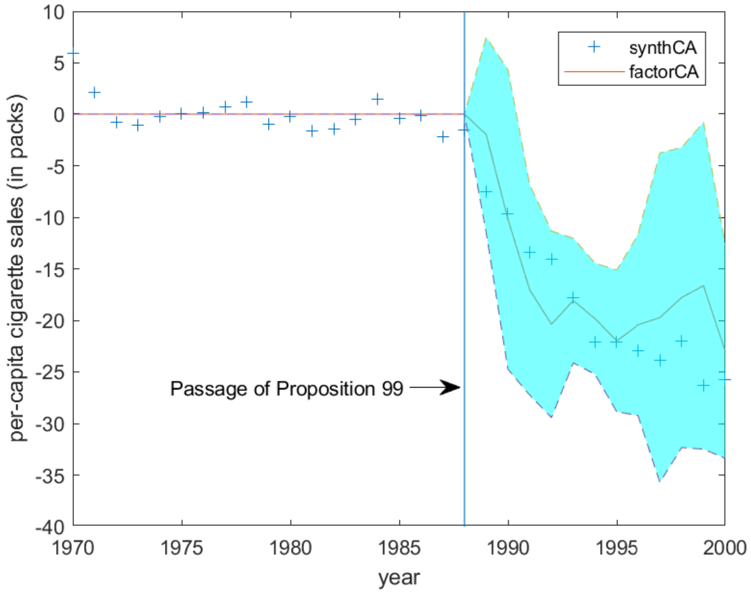

Figure 4: The 95% confidence intervals for factor causal estimates

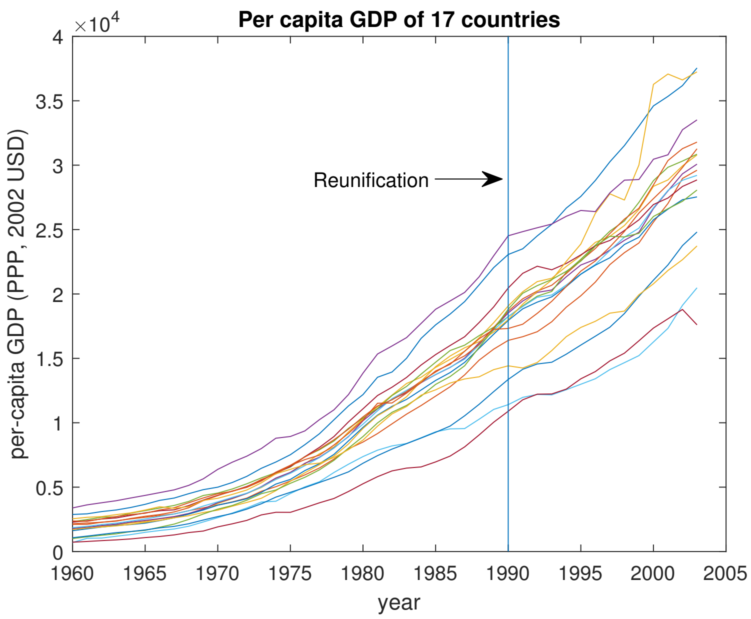

Figure 5: Time series plot of per capital GDP

strong comovement but not necessarily parallel trend. The vertical line represents year 1990 and the treatment periods are from from $T_0 = 1991$ to $T = 2003$.

We estimate the causal factor model. The factor causal effects remain to be similar to the synthetic control estimates as showcased by Figure 6 and 7.

We then regress West Germany's observed outcome on the two factors using the whole sample. The Quandt Likelihood Ratio test for a structural break at an unknown point with 15 percent trimming yields a p-value of 0.0000, with the maximum F-

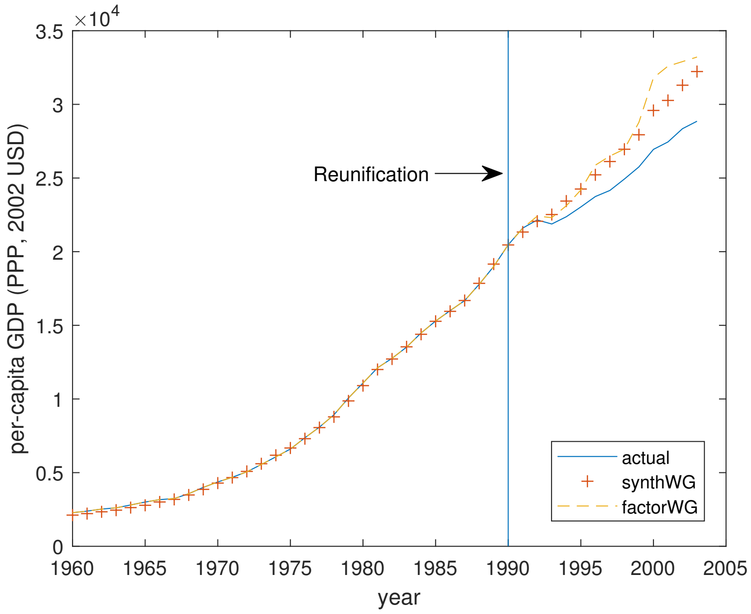

Figure 6: Causal factor model vs. synthetic control: the counterfactual West Germany

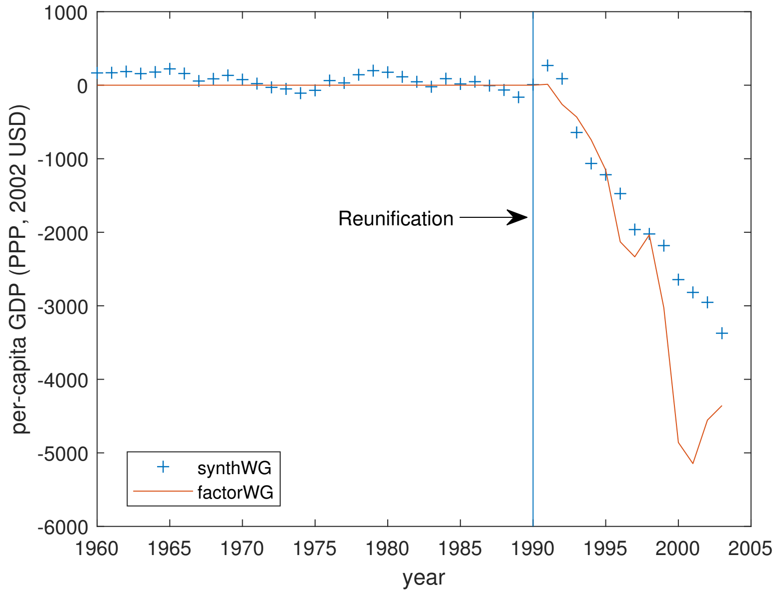

Figure 7: Causal factor model vs. synthetic control: the causal effects

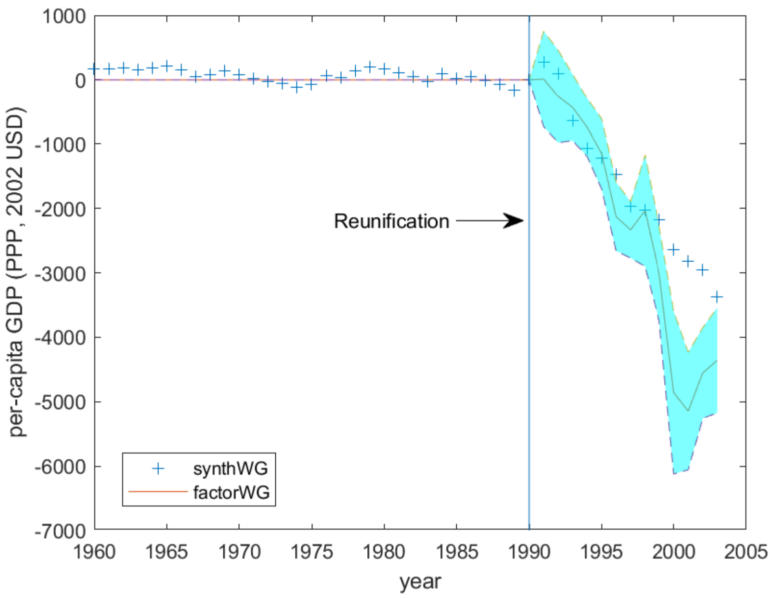

Figure 8: The 95% confidence intervals for factor causal estimates

statistic at observation 1993. In the meantime, the Chow-test for a structural break at observation 1991 yields an F-statistic of 62.45 with p-value 0.0000.

In Figure 8, we also provide the confidence intervals of the causal effects based on Section 5. The confidence intervals show that our causal estimates are mostly significant at 5% level.

Conclusion

Our study has shown that using factor models to model potential outcomes under a panel setting is a promising approach for causal analysis. By employing a causal factor model, we can explore the source of the causal effect, whether it is due to structural breaks in factor loadings, factors, or covariates' coefficients. Through two empirical examples, we have demonstrated the similarities and differences between our approach and the synthetic control method.

References

A. Abadie and J. Gardeazabal. The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93(1):113–132, 2003.

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*, 105(490):493–505, 2010.

A. Abadie, A. Diamond, and J. Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015.

J. Bai. Inferential theory for factor models of large dimensions. *Econometrica*, 71(1): 135–172, 2003.

J. Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.

J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

J. Bai and S. Ng. Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, 74(4):1133–1150, 2006.

J. Bai and S. Ng. Principal components estimation and identification of static factors. *Journal of Econometrics*, 176(1):18–29, 2013.

J. Bai and S. Ng. Matrix completion, counterfactuals, and factor analysis of missing data. *Journal of the American Statistical Association*, 116(536):1746–1763, 2021.

B. Callaway and S. Karami. Treatment effects in interactive fixed effects models with a small number of time periods. *Journal of Econometrics*, 233(1):184–208, 2023.

D. Card and AB. Krueger. Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *The American Economic Review*, 84(4):772–793, 1994.

N. Doudchenko and G. Imbens. Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. NBER Working Paper 22791, 2016.

L. Gobillon and T. Magnac. Regional policy evaluation: Interactive fixed effects and synthetic controls. *Review of Economics and Statistics*, 98(3):535–551, 2016.

C. Hsiao, HS. Ching, and SK. 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*, 27(5):705–740, 2012.

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

Appendix

The large sample theory for the causal estimator (4) is based on Assumptions A1-A5 below. Let $0 < M < \infty$ denote a generic constant, not depending on n and T.

Assumption A1. The factors satisfy $E\left\|f_t\right\| \leq M$ and $\frac{1}{T}\sum_{t=1}^{T} f_t f_t' \xrightarrow{p} \Sigma_f > 0$.

Assumption A2. The factor loadings satisfy $E\left\|\lambda_i(0)\right\|^4 \leq M < \infty$ and

$$\frac{1}{n-n_0}\sum_{i=n_0+1}^{n} \lambda_i(0)\lambda_i(0)' \xrightarrow{p} \Sigma_\lambda > 0.$$

Assumption A3. The error term ε_{it} is independent over i. In addition,

1. $E(\varepsilon_{it}) = 0$, $E\left(\varepsilon_{it}^8\right) \leq M$.

2. $E(\varepsilon_{it}\varepsilon_{is}) = \sigma_{i,ts}$ and $\left|\sigma_{i,ts}\right| \leq \bar{\sigma}_i$ for all (t,s), and $\left|\sigma_{i,ts}\right| \leq c_{ts}$ for all i such that $\frac{1}{n-n_0}\sum_{i=n_0+1}^{n}\bar{\sigma}_i \leq M$, $\frac{1}{T}\sum_{t,s=1}^{T} c_{ts} \leq M$ and

$$\frac{1}{(n-n_0)T}\sum_{i,t,s}\left|\sigma_{i,ts}\right| \leq M.$$

3. For all (t,s), $E\left|\frac{1}{\sqrt{n-n_0}}\sum_{i=n_0+1}^{n}\left[\varepsilon_{it}\varepsilon_{is} - E(\varepsilon_{it}\varepsilon_{is})\right]\right|^4 \leq M$.

4. For all t, $\frac{1}{\sqrt{n-n_0}}\sum_{i=n_0}^{n}\lambda_i(0)\varepsilon_{it} \xrightarrow{d} N(0,\Gamma_t)$, where

$$\Gamma_t = \lim_{n-n_0\to\infty}\frac{1}{n-n_0}\sum_{i=n_0}^{n}\sum_{j=n_0}^{n} E\left(\lambda_i\lambda_j' e_{it} e_{jt}\right).$$

Assumption A4. $E\left(\varepsilon_{it}|\lambda_j(d), f_s;\ d = 0,1, j = 1,...,n,\ s = 1,...,T\right) = 0$.

Assumption A5. For each $i \leq n_0$, define $z_{it} = \left[f_t', X_{it}'\right]'$, then $E\left\|z_{it}\right\|^4 \leq M$, $E(\varepsilon_{it}|z_{it}) = 0$, and $\{z_{it}, \varepsilon_{it},\ i \leq n_0\}$ is independent of $\left\{\varepsilon_{js},\ j > n_0\right\}$ for all (t,s). In addition,

1. $\frac{1}{T}\sum_{t=1}^{T} z_{it}z_{it}' \xrightarrow{p} \Sigma_i > 0$;

2. $\frac{1}{\sqrt{T_0-1}}\sum_{t=1}^{T_0-1} z_{it}\varepsilon_{it} \xrightarrow{d} N(0,\Sigma_{i0})$ and $\frac{1}{\sqrt{T-T_0+1}}\sum_{t=T_0}^{T} z_{it}\varepsilon_{it} \xrightarrow{d} N(0,\Sigma_{i1})$, where $\Sigma_{i0} = \text{plim}\frac{1}{T_0-1}\sum_{t=1}^{T_0}\left(z_{it}z_{it}'\varepsilon_{it}^2\right) > 0$ and $\Sigma_{i1} = \text{plim}\frac{1}{T-T_0+1}\sum_{t=T_0}^{T}\left(z_{it}z_{it}'\varepsilon_{it}^2\right) > 0$.

Proof of Proposition 1. Proposition 1 is implied by Theorem 3 in Bai and Ng (2006). The main extra step under our setup is to derive $var\left(\hat{\delta}_i\right)$. Note that

$$var\left(\hat{\delta}_i\right) = var\left(\hat{\delta}_i(1)\right) + var\left(\hat{\delta}_i(0)\right) - 2\cdot cov\left(\hat{\delta}_i(1), \hat{\delta}_i(0)\right).$$

As $\hat{\delta}_i(d)$, $d = 0, 1$, are separately estimated from two subsamples, we need to exploit the covariance structure across the two subsamples. From (A.1) in Bai and Ng (2006),

$$\sqrt{T(d)}\left[\hat{\delta}_i(d) - \delta_i(d)\right] = \left(\frac{1}{T(d)}\hat{z}_i(d)'\hat{z}_i(d)\right)^{-1}\frac{1}{\sqrt{T(d)}}\hat{z}_i(d)'\varepsilon_i(d) + o_p(1), \; d = 0, 1,$$

where $T(d)$ denotes the time length of the sample under treatment status $d = 0$ or 1:

$$T(0) = T_0 - 1, \; T(1) = T - T_0 + 1,$$

and

$$\hat{z}_i(0) = \begin{bmatrix} \left[\hat{f}_1', X_{i1}'\right]' \\ \vdots \\ \left[\hat{f}_{T_0-1}', X_{i,T_0-1}'\right]' \end{bmatrix}, \; \hat{z}_i(1) = \begin{bmatrix} \left[\hat{f}_{T_0}', X_{iT_0}'\right]' \\ \vdots \\ \left[\hat{f}_T', X_{i,T}'\right]' \end{bmatrix}.$$

The error term $\varepsilon_i(d)$ denotes the idiosyncratic errors for the corresponding subsamples. From Assumptions A3 and A5, $\sqrt{T - T_0 + 1}\hat{\delta}_i(1)$ and $\sqrt{T_0 - 1}\hat{\delta}_i(0)$ are asymptotically uncorrelated. As a result

$$var\left(\hat{\delta}_i\right) = var\left(\hat{\delta}_i(1)\right) + var\left(\hat{\delta}_i(0)\right) + o_p\left(\frac{1}{\sqrt{T - T_0}\sqrt{T_0}}\right).$$

The corresponding $var\left(\hat{\delta}_i(d)\right)$ is readily available from Bai and Ng (2006):

$$var\left(\hat{\delta}_i(d)\right)$$
$$= \frac{1}{T(d)}\left(\frac{1}{T(d)}\hat{z}_i(d)'\hat{z}_i(d)\right)^{-1}\left(\frac{1}{T(d)}\sum_{t=1}^{T(d)}\hat{\varepsilon}_{it}(d)^2\hat{z}_{it}(d)\hat{z}_{it}(d)'\right)\left(\frac{1}{T(d)}\hat{z}_i(d)'\hat{z}_i(d)\right)^{-1}.$$

Q.E.D.

Proof of Proposition 2. For notational simplicity, define $C_{it}(d) = \lambda_i(d)' f_t(d)$, $d = 0, 1$, $n_1 = n - n_0$, $T_1 = T - T_0 + 1$, $\delta_{n_0 T_1} = min\left\{\sqrt{n_0}, \sqrt{T_1}\right\}$.

Bai (2003) provides the asymptotic expansion of $\hat{C}_{it} - C_{it}$ for $i \leq n_0$, $t \geq T_0$:

$$\hat{C}_{it}(1) - C_{it}(1) = \lambda_i(1)'\left(\frac{\sum_{k=1}^{n_0}\lambda_k(1)\lambda_k(1)'}{n_0}\right)^{-1}\frac{1}{n_0}\sum_{k=1}^{n_0}\lambda_k(1)\varepsilon_{kt}$$
$$+ f_t(1)'\left(\frac{\sum_{s=T_0}^{T} f_s(1) f_s(1)'}{T_1}\right)^{-1}\frac{1}{T_1}\sum_{s=T_0}^{T} f_s(1)\varepsilon_{is}$$
$$+ O_p\left(\frac{1}{\delta_{n_0 T_1}^2}\right)$$

$$
\begin{aligned}
&\equiv M_A \sum_{k=1}^{n_0} \lambda_k(1)\,\varepsilon_{kt} + M_B \sum_{s=T_0}^{T} f_s(1)\,\varepsilon_{is} + O_p\left(\frac{1}{\delta^2_{n_0T_1}}\right) \\
&= A + B + O_p\left(\frac{1}{\delta^2_{n_0T_1}}\right).
\end{aligned}
$$

The asymptotic expansion for $\hat{f}_t(0)$ is based on the full sample principal component analysis of untreated units. Note that the product $\lambda_i(d)' f_t(d)$ does not depend on how the factors and factor loadings are rotated in the estimation. We will maintain the assumptions PC1 in Bai and Ng (2013) that the same identification constraints hold for both the true and the estimated factors and factor loadings. This means that one can simply take the rotation matrix as an identity matrix in terms of asymptotic representations. Assume $\sqrt{n}/T \to 0$, we have

$$
\sqrt{n_1}\left(\hat{f}_t(0) - f_t(0)\right) = \left(\frac{\sum_{k=n_0+1}^{n} \lambda_k(0)\,\lambda_k(0)'}{n_1}\right)^{-1} \frac{1}{\sqrt{n_1}} \sum_{k=n_0+1}^{n} \lambda_k(0)\,\varepsilon_{kt} + o_p(1).
$$

Then $\hat{\lambda}_i(0)$ is the OLS estimator from regression y_{it} on $\hat{f}_t(0)$ for $t < T_0$, whose asymptotic expansion is based on Bai and Ng (2006):

$$
\sqrt{T_0 - 1}\left(\hat{\lambda}_i(0) - \lambda_i(0)\right) = \left(\frac{\sum_{s=1}^{T_0-1} f_s(0)\,f_s(0)'}{T_0 - 1}\right)^{-1} \frac{1}{\sqrt{T_0 - 1}} \sum_{s=1}^{T_0-1} f_s(0)\,\varepsilon_{is} + o_p(1).
$$

Denote $\delta_{nT} = min\left\{\sqrt{n}, \sqrt{T}\right\}$. Under the assumption $n_0/n \to c \in (0,1)$, $T_0/T \to b \in (0,1)$, we obtain (for $i \le n_0,\ t \ge T_0$):

$$
\begin{aligned}
\hat{C}_{it}(0) - C_{it}(0) &= \left[\hat{\lambda}_i(0) - \lambda_i(0)\right]' \hat{f}_t(0) + \lambda_i(0)' \left[\hat{f}_t(0) - f_t(0)\right] \\
&= \frac{1}{T_0 - 1} \hat{f}_t(0)' \left(\frac{\sum_{s=1}^{T_0-1} f_s(0)\,f_s(0)'}{T_0 - 1}\right)^{-1} \sum_{s=1}^{T_0-1} f_s(0)\,\varepsilon_{is} \\
&\quad + \frac{1}{n_1} \lambda_i(0)' \left(\frac{\sum_{k=n_0+1}^{n} \lambda_k(0)\,\lambda_k(0)'}{n_1}\right)^{-1} \sum_{k=n_0+1}^{n} \lambda_k(0)\,\varepsilon_{kt} \\
&\quad + O_p\left(\frac{1}{\delta^2_{nT}}\right) \\
&\equiv C + D + O_p\left(\frac{1}{\delta^2_{nT}}\right)
\end{aligned}
$$

where the second equality uses the asymptotic expansion for $\hat{\lambda}_i(0) - \lambda_i(0)$ and that of

$\hat{f}_t(0) - f_t(0)$ given earlier. In sum, for $i \le n_0,\ t \ge T_0$

$$\begin{aligned}\hat{\tau}_{it} - \tau_{it} &= \hat{C}_{it}(1) - \hat{C}_{it}(0) - [C_{it}(1) - C_{it}(0)] \\ &= \hat{C}_{it}(1) - C_{it}(1) - \left[\hat{C}_{it}(0) - C_{it}(0)\right] \\ &= A + B - C - D + O_p\left(\frac{1}{\delta_{nT}^2}\right).\end{aligned}$$

Using the asymptotic expansion, let V_{it} denote the variance of $A + B - C - D$, then

$$V_{it}^{-1/2}(\hat{\tau}_{it} - \tau_{it}) \xrightarrow{d} N(0,1).$$

Q.E.D.

We now provide an estimate for V_{it} in Proposition 2. Assume ε_{it} are uncorrelated over i and t, then it is clear that A, B, C, D are mutually uncorrelated. So the variance of $\hat{\tau}_{it}$ is given by

$$var(\hat{\tau}_{it} - \tau_{it}) = var(A) + var(B) + var(C) + var(D) + O\left(\frac{1}{\delta_{nT}^4}\right).$$

We estimate each component by

$$\begin{aligned}\widehat{var(A)} &= \hat{M}_A \sum_{k=1}^{n_0} \hat{\lambda}_k(1)\hat{\lambda}_k(1)' \hat{\varepsilon}_{kt}^2 \hat{M}_A', \\ \widehat{var(B)} &= \hat{M}_B \sum_{s=T_0}^{T} \hat{f}_s(1)\hat{f}_s(1)' \hat{\varepsilon}_{is}^2 \hat{M}_B, \\ \widehat{var(C)} &= \frac{1}{(T_0-1)^2}\hat{f}_t(0)' \hat{W}^{-1}\left[\sum_{s=1}^{T_0-1} \hat{f}_s(0)\hat{f}_s(0)' \hat{\varepsilon_{is}}^2\right]\hat{W}^{-1}\hat{f}_t(0), \\ \widehat{var(D)} &= \frac{1}{n_1^2}\hat{\lambda}_i(0)' \hat{V}^{-1}\left[\sum_{k=n_0+1}^{n} \hat{\lambda}_k(0)\hat{\lambda}_k(0)' \hat{\varepsilon}_{kt}^2\right]\hat{V}^{-1}\hat{\lambda}_i(0),\end{aligned}$$

where $\hat{W} = \frac{1}{T_0-1}\sum_{s=1}^{T_0-1} \hat{f}_s(0)\hat{f}_s(0)'$ and $\hat{V}$ denotes the diagonal matrix consisting of the first r eigenvalues of $\frac{1}{n_1 T}\sum_{i=n_0+1}^{n} y_i y_i'$ in decreasing order, with y_i being the $T \times 1$ observed outcomes for individual i.