Supplementary Materials for

Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models

Yiqing Xu

University of California, San Diego

Table of Contents

A Technical Details . . . A-2
- A.1 A Directed Acyclic Graph (DAG) . . . A-2
- A.2 Technical Assumptions . . . A-3
- A.3 Estimating an Interactive Fixed-effect Model . . . A-4
- A.4 Estimation Procedure for an Extended Model . . . A-5
- A.5 Proofs of Main Results . . . A-6

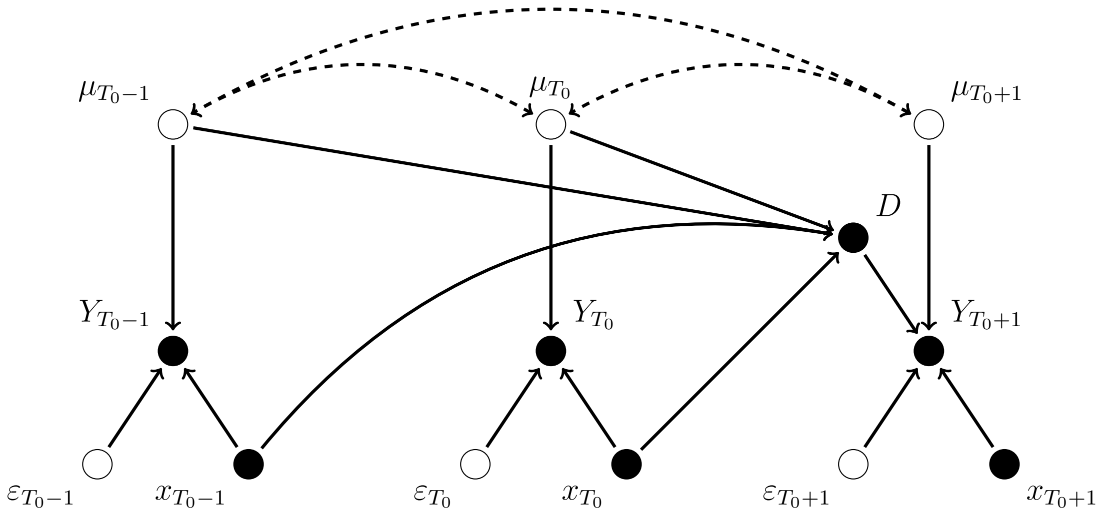
B Simulation Results . . . A-12
- B.1 Additional Results on The Simulated Example . . . A-12
- B.2 Finite Sample Properties . . . A-14
- B.3 Comparison with the DID Estimator . . . A-15
- B.4 Comparison with the IFE Estimator . . . A-16
- B.5 Comparison with the Synthetic Matching Method . . . A-17
- B.6 Model Selection . . . A-19

C Additional Results on EDR Laws and Turnout . . . A-21

A Technical Details

A.1 A Directed Acyclic Graph (DAG)

Figure A1. A DAG Illustration



Note: Unit indices are dropped for simplicity. Vector μ_t represents unobserved time-varying confounders. If Assumption 1 holds, μ_t (or μ_{it}) can be expressed as $\lambda_i' f_t$. We allow D_i to be correlated with $x_{is,s<t}$ and $\mu_{is,s<t}$. In fact, we also allow it to be correlated with $x_{js,s<t}$ and $\mu_{js,s<t}$ when $j \neq i$.

A.2 Technical Assumptions

Assumptions 3–5 are shown below.

Assumption 3 Weak serial dependence of the error terms:

1. $\mathbb{E}(\varepsilon_{it}\varepsilon_{is}) = \sigma_{i,ts}$, $|\sigma_{i,ts}| \leq \bar{\sigma}_i$ for all (t, s) such that $\frac{1}{N}\sum_{i=1}^{N}\bar{\sigma}_i \leq M$.
2. For every (t, s), $\mathbb{E}|N^{-1/2}\sum_{i=1}^{N}[\varepsilon_{is}\varepsilon_{it} - \mathbb{E}(\varepsilon_{is}\varepsilon_{it})]|^4 \leq M$.
3. $\frac{1}{T^2N}\sum_{t,s,u,v}\sum_{i,j}|cov(\varepsilon_{it}\varepsilon_{is}, \varepsilon_{ju}\varepsilon_{jv})| \leq M$ and $\frac{1}{TN^2}\sum_{t,s}\sum_{i,j,k,l}|cov(\varepsilon_{it}\varepsilon_{jt}, \varepsilon_{ks}\varepsilon_{ls})| \leq M$.
4. $\mathbb{E}(\varepsilon_{it}\varepsilon_{js}) = 0$, for all $i \neq j$, (t, s).

Assumption 4 Regularity conditions:

1. $\mathbb{E}|\varepsilon_{it}|^8 \leq M$.
2. $\mathbb{E}\|x_{it}\|^4 \leq M$: Let $\mathcal{F} = \{F : F'F/T = I\}$. We assume $\inf_{F\in\mathcal{F}} D(F) > 0$, in which $D(F) = \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}} S_i'S_i$, where $S_i = M_F X_i - \frac{1}{N_{co}}\sum_{k=1}^{N} M_F X_k a_{ik}$ and $a_{ik} = \lambda_i'(\Lambda_{co}'\Lambda_{co})^{-1}\lambda_k$.
3. $\mathbb{E}\|f_t\|^4 \leq M < \infty$ and $\frac{1}{T}\sum_{t=1}^{T} f_t f_t' \xrightarrow{p} \Sigma_F$ for some $r \times r$ positive definite matrix Σ_F, as $T_0 \to \infty$.
4. $\mathbb{E}\|\lambda_i\|^4 \leq M < \infty$ and $\Lambda_{co}'\Lambda_{co}/N_{co} \xrightarrow{p} \Sigma_N$ for some $r \times r$ positive definite matrix Σ_N, as $N_{co} \to \infty$.

Assumption 5 The error terms are cross-sectionally independent and homoscedastic.

1. $\varepsilon_{it} \perp\!\!\!\perp \varepsilon_{js}$ for all $j \neq i$, (t, s).
2. $\mathbb{E}(\varepsilon_{it}\varepsilon_{is}) = \sigma_{ts} \leq M$, for all (t, s).

A.3 Estimating an Interactive Fixed-effect Model

As in Equation (1), I assume that the control units follow an interactive fixed-effect model:

$$Y_{co} = X_{co}\beta + F\Lambda'_{co} + \varepsilon_{co},$$

The least square objective function is

$$SSR(\beta, F, \Lambda_{co}) = \sum_{i=1}^{N_{co}} (Y_i - X_i\beta - F\lambda_i)'(Y_i - X_i\beta - F\lambda_i).$$

The goal is to estimate β, F, and Λ_{co} by minimizing the SSR subject to the following constraints:

$$F'F/T = I_r \text{ and } \Lambda'_{co}\Lambda_{co} = \text{diagonal}.$$

A unique solution $(\hat{\beta}, \hat{F}, \hat{\Lambda}_{co})$ to this problem exists. To find the solution, Bai (2009) proposed an iteration scheme that can lead to the unique solution starting from some initial value of β (for instance, the least-square dummy-variable (LSDV) estimates) or F. In each iteration, given F and Λ_{co}, the algorithm computes β:

$$\hat{\beta}(F, \Lambda) = \left(\sum_{i=1}^{N_{co}} X'_i X_i \right)^{-1} \sum_{i=1}^{N_{co}} X'_i (Y_i - F\lambda_i),$$

and given β, it computes F and Λ_{co} from a pure factor model $(Y_i - X_i\beta) = F\lambda_i + \varepsilon_i$:

$$[\tfrac{1}{N_{co}T} \textstyle\sum_{i=1}^{N_{co}} (Y_i - X_i\beta)(Y_i - X_i\beta)'] \hat{F} = \hat{F} V_{N_{co}T},$$
$$\hat{\Lambda}_{co} = \tfrac{1}{T}(Y - X\beta)'\hat{F},$$

in which $V_{N_{co}T}$ is a diagonal matrix that consists for the first r largest eigenvalues of the $(N_{co} \times N_{co})$ matrix $\frac{1}{N_{co}T} \sum_{i=1}^{N_{co}} (Y_i - X_i\beta)(Y_i - X_i\beta)'$ and $V_{N_{co}T} = \frac{1}{N_{co}} \hat{\Lambda}'_{co}\hat{\Lambda}_{co}$.

A.4 Estimation Procedure for an Extended Model

Without loss of generality, we re-write Equation (2) as

$$Y_{it} = \delta_{it} D_{it} + x'_{it}\beta + \gamma'_i l_t + z'_i \theta_t + \lambda'_i f_t + \alpha_i + \xi_t + \mu + \varepsilon_{it},$$

in which μ is the mean of control group outcomes, which allows us to impose two restrictions: $\sum_{i=1}^{N_{co}} \alpha_i = 0$ and $\sum_{i=1}^{N_{co}} \xi_i = 0$. As before, we use three steps to impute the counterfactuals for treated units. It can be written as

$$Y_i = \delta_i \circ D_i + X_i\beta + L\gamma_i + \Theta z_i + F\lambda_i + \alpha_i \mathbf{1}_T + \Xi + \mu \mathbf{1}_T + \varepsilon_i,$$

in which $L = [l_1, l_2, \cdots, l_T]'$, a $(T \times q)$ matrix; $\Theta = [\theta_1, \theta_2, \cdots, \theta_T]'$, a $(T \times m)$ matrix; and $\Xi = [\xi_1, \xi_2, \cdots, \xi_T]'$, a $(T \times 1)$ vector. In the first step, we estimate an extended IFE model using only the control group data and obtain $\hat{\beta}$,$\hat{F}$,$\hat{\Lambda}_{co}$, $\hat{\Xi}$, $\hat{\Theta}$, $\hat{\gamma}_i$, and $\hat{\alpha}_i$ (for all $i \in \mathcal{C}$) and $\hat{\mu}$:

$$\textbf{Step 1.} \quad \left(\hat{\beta}, \hat{F}, \hat{\Theta}, \hat{\Xi}, \hat{\Lambda}_{co}, \{\hat{\gamma}_i\}, \{\hat{\alpha}_i\}, \hat{\mu}\right) = \underset{\tilde{\beta}, \tilde{F}, \tilde{\Theta}, \tilde{\Xi}, \tilde{\Lambda}_{co}, \{\tilde{\gamma}_i\}, \{\tilde{\alpha}_i\}, \tilde{\mu}}{\operatorname{argmin}} \sum_{i \in \mathcal{C}} \tilde{e}'_i \tilde{e}_i.$$

in which $\tilde{e}_i = Y_i - X_i\tilde{\beta} - L\tilde{\gamma}_i - \tilde{\Theta} z_i - \tilde{F}\tilde{\lambda}_i - \tilde{\alpha}_i \mathbf{1}_T - \tilde{\Xi} - \tilde{\mu}\mathbf{1}_T$. The details of the estimation strategy can be found in Bai (2009) Sections 8 and 10.

The second step estimates factor loadings, as well as additive unit fixed effects, for each treated unit by minimizing the mean squared error of the treated units in the pre-treatment period:

$$\begin{aligned} \textbf{Step 2.} \quad \left(\hat{\gamma}_i \ \hat{\lambda}_i \ \hat{\alpha}_i\right)' &= \underset{(\tilde{\gamma}_i \ \tilde{\lambda}_i \ \tilde{\alpha}_i)'}{\operatorname{argmin}} \, \underline{e}'_i \underline{e}_i \\ &= (\hat{G}^{0\prime}\hat{G}^0)^{-1}\hat{G}^{0\prime}(Y_i^0 - X_i^0\hat{\beta} - \hat{\Theta}^0 z_i - \hat{\Xi}^0 - \hat{\mu}\mathbf{1}_{T_0}), \qquad i \in \mathcal{T}. \end{aligned}$$

in which $\underline{e}_i = Y_i^0 - X_i^0\hat{\beta} - L^0\tilde{\gamma}_i - \hat{\Theta}^0 z_i - \hat{F}^0\tilde{\lambda}_i - \tilde{\alpha}_i - \hat{\xi}_t^0 - \hat{\mu}$; $\hat{\beta}$, $\hat{F}$, $\hat{\Theta}$, $\hat{\Xi}$ and $\hat{\mu}$ are from the first step estimation; the superscripts "0"s denote the pre-treatment period; and $\hat{G}^0 = (L^0 \ \hat{F}^0 \ \mathbf{1}_{T_0})$ is $(L^0 \ \hat{F}^0)$ augmented with a column of ones, a $T_0 \times (q + r + 1)$ matrix.

In the third step, we calculate the counterfactual based on $\hat{\beta}$, $\hat{F}$, $\hat{\lambda}_i$, $\hat{\alpha}_i$, $\hat{\xi}_t$, and $\hat{\mu}$:

$$\textbf{Step 3.} \qquad \hat{Y}_{it}(0) = x'_{it}\hat{\beta} + \hat{\gamma}'_i l_t + z'_i \hat{\theta}_t + \hat{\lambda}'_i \hat{f}_t + \hat{\alpha}_i + \hat{\xi}_t + \hat{\mu}, \qquad \forall i \in \mathcal{T}.$$

A.5 Proofs of Main Results

We use the Frobenius norm throughout this paper, i.e., for any vector or matrix M, its norm is defined as $\|M\| = \sqrt{\text{tr}(M'M)}$. I establish four lemmas before getting to the main results.

Lemma 1 *(i)* $T^{-1/2}\|\hat{F}^0\| = O_p(1)$; *(ii)* $T\|(\hat{F}^{0\prime}\hat{F}^0)^{-1}\| = O_p(1)$.

Proof: (i). Because $\text{tr}(\hat{F}'\hat{F}/T) = r$,

$$T^{-1/2}\|\hat{F}^0\| = T^{-1/2}\sqrt{\text{tr}(\hat{F}^{0\prime}\hat{F}^0)} \le T^{-1/2}\sqrt{\text{tr}(\hat{F}'\hat{F})} = \sqrt{r}.$$

(ii). Denote $Q = \sum_{s=T_0+1}^{T} \hat{f}_s\hat{f}_s'$, a symmetric and positive definite $(r \times r)$ matrix. Because $\|\hat{f}_t\hat{f}_t'\| = O_p(1)$ and there are only q_i items in the summation, $\|Q\| = O_p(1)$. Since $\hat{F}^{0\prime}\hat{F}^0 = \hat{F}'\hat{F} - Q = T \cdot I_r - Q$,

$$(\hat{F}^{0\prime}\hat{F}^0)^{-1} = \frac{1}{T}I_r + (I - \frac{1}{T}Q)^{-1}\frac{1}{T^2}Q$$

Since Q is positive definite, $\|(\hat{F}^{0\prime}\hat{F}^0)^{-1}\|$ is strictly decreasing in T and is $O_p(T^{-1})$.

Lemma 2 $\|\hat{\beta} - \beta\| = O_p(N_{co}^{-1}) + O_p(T^{-1}) + o_p\left((N_{co}T)^{-1/2}\right)$.

Proof: Bai (2009) shows that under Assumptions 3 and 4 and when $T/N^2 \to 0$:

$$\begin{aligned}\hat{\beta} - \beta =& D(\hat{F})^{-1}\frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}[X_i'M_F - \frac{1}{N_{co}}\sum_{k=1}^{N_{co}} a_{ik}X_k'M_F]\varepsilon_i \\ &+ \frac{1}{N_{co}}\xi + \frac{1}{T}\zeta + \frac{1}{\sqrt{N_{co}T}}o_p(1),\end{aligned} \tag{1}$$

$$\begin{aligned}\text{where } D(\hat{F}) =& \frac{1}{N_{co}T}\sum_{i}^{N_{co}} Z_i'Z_i, \quad Z_i = M_FX_i - \frac{1}{N_{co}}\sum_{k=1}^{N_{co}} M_FX_ka_{ik}, \\ \xi =& -D(\hat{F})^{-1}\frac{1}{N_{co}}\sum_{i=1}^{N_{co}}\sum_{k=1}^{N_{co}}\frac{(X_i - V_i)'F}{T}\left(\frac{F'F}{T}\right)^{-1}\left(\frac{\Lambda_{co}'\Lambda_{co}}{N_{co}}\right)^{-1}\lambda_k\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}\varepsilon_{kt}\right) = O_p(1), \\ \zeta =& -D(\hat{F})^{-1}\frac{1}{NT}\sum_{i=1}^{N_{co}} X_i'M_{\hat{F}}\Omega\hat{F}\left(\frac{F'\hat{F}}{T}\right)^{-1}\left(\frac{\Lambda_{co}'\Lambda_{co}}{N_{co}}\right)^{-1}\lambda_i = O_p(1),\end{aligned}$$

and $a_{ik} = \lambda_i'(\Lambda_{co}'\Lambda_{co}/N_{co})^{-1}\lambda_k, V_i = \frac{1}{N_{co}}\sum_{k=1}^{N_{co}} a_{ik}X_k, \Omega = \frac{1}{N_{co}}\sum_{k=1}^{N_{co}} E(\varepsilon_k\varepsilon_k')$. Therefore, $\hat{\beta}$ is an asymptotically unbiased estimator for β when both T and N_{co} are large and

$$\|\hat{\beta} - \beta\| = O_p(N_{co}^{-1}) + O_p(T^{-1}) + o_p\left((N_{co}T)^{-1/2}\right).$$

Lemma 3 *Denote* $H = \left(\frac{\Lambda_{co}'\Lambda_{co}}{N_{co}}\right)\left(\frac{F'\hat{F}}{T}\right)V_{N_{co}T}^{-1}$.

(i). $\|f_t - H^{-1}\hat{f}_t\| = O_p(N_{co}^{-1/2}) + O_p(T^{-1/2})$;

(ii). $\|f_t' - \hat{f}_t{}'(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{F}^{0\prime}F^0\| = O_p(N_{co}^{-1/2}) + O_p(T^{-1/2})$.

Proof: (i). The main logic of this proof follows Bai (2009) Proposition A.1 (p. 1266). Because

$$\left[\frac{1}{N_{co}T}\sum_{i}^{N_{co}}(Y_i - X_i)(Y_i - X_i)'\right]\hat{F} = \hat{F}V_{N_{co}T}$$

and $Y_i - X_i\hat{\beta} = X_i(\beta - \hat{\beta}) + F\lambda_i + \varepsilon_i$, by expanding the terms on the left-hand side, we have:

$$\begin{aligned}\hat{F}V_{N_{co}T} =& \frac{1}{N_{co}T}\sum_{i}^{N_{co}}X_i(\beta-\hat{\beta})(\beta-\hat{\beta})'X_i'\hat{F} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}X_i(\beta-\hat{\beta})\lambda_i'F'\hat{F} \\ &+ \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}X_i(\beta-\hat{\beta})\varepsilon'\hat{F} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}F\lambda_i(\beta-\hat{\beta})'X_i'\hat{F} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}\varepsilon_i(\beta-\hat{\beta})'X_i'\hat{F} \\ &+ \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}F\lambda_i\varepsilon_i'\hat{F} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}\varepsilon_i\lambda_i'F'\hat{F} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}\varepsilon_i\varepsilon_i'\hat{F} + F\left(\frac{\Lambda_{co}'\Lambda_{co}}{N_{co}}\right)\left(\frac{F'\hat{F}}{T}\right).\end{aligned}$$

with the last term on the right-hand side equal to $\frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}F\lambda_i\lambda_i'F'\hat{F}$. Denote $G = \left(\frac{F'\hat{F}}{T}\right)^{-1}\left(\frac{\Lambda_{co}'\Lambda_{co}}{N_{co}}\right)^{-1}$. After re-arranging the terms and focusing on period t, we have:

$$\begin{aligned}H^{-1}\hat{f}_t - f_t =& \frac{1}{N_{co}T}\sum_{i}^{N_{co}}G\hat{F}'X_i'(\beta-\hat{\beta})(\beta-\hat{\beta})'x_{it} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'F\lambda_i(\beta-\hat{\beta})'x_{it} \\ &+ \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'\varepsilon_i(\beta-\hat{\beta})'x_{it} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'X_i'(\beta-\hat{\beta})\lambda_i'f_t \\ &+ \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'X_i(\beta-\hat{\beta})\varepsilon_{it} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'\varepsilon_i\lambda_i'f_t \\ &+ \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'F\lambda_i\varepsilon_{it} + \frac{1}{N_{co}T}\sum_{i=1}^{N_{co}}G\hat{F}'\varepsilon_i\varepsilon_{it} \\ &= a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8,\end{aligned}$$

The proof of $\|(\frac{F'\hat{F}}{T})^{-1}\| = O_p(1)$ is provided in Bai (2003) Proposition 1. Assumption 4 implies $\|(\frac{\Lambda_{co}'\Lambda_{co}}{N_{co}})^{-1}\| = O_p(1)$, therefore, $\|G\| = O_p(1)$. Also from Assumption 4, we know that $\|x_{it}\| = O_p(1)$, $T^{-1/2}\|X_i\| = O_p(1)$, $N_{co}^{-1/2}\|\Lambda_{co}\| = O_p(1)$. Together with the facts that $T^{-1/2}\|F\| = O_p(1)$ and $T^{-1/2}\|\hat{F}\| = \sqrt{r}$, we have:

$$\|a_1\| \leq \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{1}{N_{co}} \sum_{i}^{N_{co}} \left(\frac{\|X_i\|}{\sqrt{T}} \|x_{it}\| \right) \|\beta - \hat{\beta}\|^2 = O_p(\|\beta - \hat{\beta}\|^2)$$

Similarly, we can show that each of a_2, a_3, a_4, and a_5 is $O_p(\beta - \hat{\beta})$.

$$\begin{aligned}
\|a_2\| &\leq \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{\|F\|}{\sqrt{T}} \frac{1}{N_{co}} \sum_{i=1}^{N_{co}} (\|\lambda_i\| \|x_{it}\|) \|\beta - \hat{\beta}\| = O_p(\|\beta - \hat{\beta}\|) \\
\|a_3\| &\leq \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{1}{N_{co}} \sum_{i=1}^{N_{co}} \left(\frac{\|\varepsilon_i\|}{\sqrt{T}} \|x_{it}\| \right) \|\beta - \hat{\beta}\| = O_p(\|\beta - \hat{\beta}\|) \\
\|a_4\| &\leq \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{1}{N_{co}} \sum_{i=1}^{N_{co}} \left(\frac{\|X_i\|}{\sqrt{T}} \|\lambda_i' f_t\| \right) \|\beta - \hat{\beta}\| = O_p(\|\beta - \hat{\beta}\|) \\
\|a_5\| &\leq \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{1}{N_{co}} \sum_{i=1}^{N_{co}} \left(\frac{\|X_i\|}{\sqrt{T}} \|\varepsilon_{it}\| \right) \|\beta - \hat{\beta}\| = O_p(\|\beta - \hat{\beta}\|)
\end{aligned}$$

Moreover, a_6 and a_7 are both $O_p(N^{-1/2})$.

$$\begin{aligned}
\|a_6\| &= \frac{1}{N_{co}T} \|G \hat{F}' \varepsilon \Lambda_{co}' f_t\| \leq \frac{1}{\sqrt{N_{co}}} \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{\|\varepsilon\|}{\sqrt{T}} \frac{\|\Lambda_{co}\|}{\sqrt{N_{co}}} \|f_t\| = O_p(N_{co}^{-1/2}) \\
\|a_7\| &\leq \frac{1}{\sqrt{N_{co}}} \|G\| \frac{\|\hat{F}\|}{\sqrt{T}} \frac{\|F\|}{\sqrt{T}} \sqrt{\frac{1}{N_{co}} \sum_{i=1}^{N_{co}} \|\lambda_i \varepsilon_{it}\|^2} = O_p(N_{co}^{-1/2})
\end{aligned}$$

Finally, Denote $\tilde{f}_t = G f_t,\ t = 1, 2, \cdots, T$,

$$\begin{aligned}
a_8 &= \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{f}_t \frac{1}{N_{co}} \sum_{i=1}^{N_{co}} \varepsilon_{is} \varepsilon_{it} \right) \\
&= \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{f}_t \frac{1}{N_{co}} \sum_{i=1}^{N_{co}} [\varepsilon_{is} \varepsilon_{it} - E(\varepsilon_{is} \varepsilon_{it})] \right) - \frac{1}{T} \sum_{s=1}^{T} \left(\tilde{f}_t E(\varepsilon_{is} \varepsilon_{it}) \right) \\
&= b_1 + b_2
\end{aligned}$$

Because $E(\varepsilon_{is}\varepsilon_{it})$ is bounded according to Assumption 4.2,

$$\|b_2\| \leq \frac{1}{\sqrt{T}} \|G\| \frac{\|F\|}{\sqrt{T}} M = O_p(T^{-1/2}).$$

On the other hand,

$$\|b_1\| \leq \frac{1}{\sqrt{N_{co}}}\|G\|\frac{\|\hat{F}\|}{\sqrt{T}}\sqrt{\frac{1}{T}\sum_{s=1}^{T}\frac{1}{N_{co}}\sum_{i=1}^{N_{co}}|\varepsilon_{it}\varepsilon_{is}-E(\varepsilon_{it}\varepsilon_{is})|^2} = O_p(N_{co}^{-1/2})$$

Therefore, $a_8 = O_p(N_{co}^{-1/2}) + O_p(T^{-1/2})$.

Because $\|\beta - \hat{\beta}\| = O_p(N_{co}^{-1}) + O_p(T^{-1}) + o_p\left((N_{co}T)^{-1/2}\right)$ according to Lemma 2,

$$\|f_t - H^{-1}\hat{f}_t\| = O_p(\|\beta - \hat{\beta}\|) + O_p(N_{co}^{-1/2}) = O_p(N_{co}^{-1/2}) + O_p(T^{-1/2}).$$

(ii). By subtracting $H^{-1}\hat{f}_t$ from $f_t - F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t$ and then adding it back, we have:

$$f_t - F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t = (f_t - H^{-1}\hat{f}_t) - (F' - H^{-1}\hat{F}')\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t$$

Because $T^{-1/2}\|F - H^{-1}\hat{F}\| = O_p(N_{co}^{-1/2}) + O_p(T^{-1/2})$ (Bai 2009, p. 1268) and $\|\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t\| = O_p(T^{-1/2})$,

$$\begin{aligned}
\|f_t - F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t\| &\leq \|f_t - H^{-1}\hat{f}_t\| + \|(F' - H^{-1}\hat{F}')\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t\| \\
&\leq \|f_t - H^{-1}\hat{f}_t\| + \|F' - H^{-1}\hat{F}'\|\|\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t\| \\
&= O_p(N_{co}^{-1/2}) + O_p(T^{-1/2})
\end{aligned}$$

It is worth noting that if $E(\varepsilon_{is}\varepsilon_{it}) = 0$ for any i and all (s,t), then

$$\|f_t - F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t\| = O_p(N_{co}^{-1/2}) + O_p(T^{-1}).$$

Lemma 4 $\|F'\hat{F}(\hat{F}'\hat{F})^{-1} - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\| = o_p(T_0^{-1})$.

Proof: Denote $A = F'\hat{F}$ and $B = F'\hat{F} - F^{0\prime}\hat{F}^0$. Both are $(r \times r)$ matrices. $\|B\| = \|\sum_{s=T_0}^{T} f_s\hat{f}_s'\| = O_p(1)$. Recall $Q = \hat{F}'\hat{F} - \hat{F}^{0\prime}\hat{F}^0$ and $\hat{F}'\hat{F}/T = I_r$.

$$\begin{aligned}
& F'\hat{F}(\hat{F}'\hat{F})^{-1} - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1} \\
=\ & \frac{1}{T}A - (A - B)\left[\frac{1}{T}I_r + (I - \frac{1}{T}Q)^{-1}\frac{1}{T^2}Q\right] \\
=\ & \frac{1}{T}B - (A - B)(I - \frac{1}{T}Q)^{-1}\frac{1}{T^2}Q
\end{aligned}$$

The second term on the right is $O_p(T^{-1})$ because $T^{-1}\|A - B\| = O_p(1)$. Therefore, $\|F'\hat{F}(\hat{F}'\hat{F})^{-1} - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\| = O_p(T^{-1})$.

Proposition 1 (Limit of Bias) *Under Assumptions 1-4,* $\mathbb{E}_\varepsilon(\widehat{ATT}_t|D, X, \Lambda, F) \to ATT_t$*, in which* $ATT_t = \frac{1}{N_{tr}} \sum_{i \in \mathcal{T}} \delta_{it}$*, for all* $t > T_0$*, as both* N_{co} *and* $T_0 \to \infty$.

Proof: Denote i as the treated unit on which the treatment effect is of interest. From $Y_{it} = x_{it}'\beta + \lambda_i' f_t + \varepsilon_{it}$ and $\hat{\lambda}_i = (\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{F}^{0\prime}(Y_i^0 - X_i^0\hat{\beta})$, we have:

$$\begin{aligned}
\hat{\delta}_{it} - \delta_{it} &= Y_{it} - \hat{Y}_{it}(0) - \delta_{it} \\
&= x_{it}'(\beta - \hat{\beta}) + (\lambda_i' f_t - \hat{\lambda}_i'\hat{f}_t) + \varepsilon_{it} \\
&= x_{it}'(\beta - \hat{\beta}) + \left\{ \lambda_i' f_t - [X_i^0(\beta - \hat{\beta}) + F^0\lambda_i + \varepsilon_i^0]'\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{f}_t \right\} + \varepsilon_{it} \\
&= \left[x_{it}' - \hat{f}_t'(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{F}^{0\prime}X_i^0 \right](\beta - \hat{\beta}) + \lambda_i' \left[f_t - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{f}_t \right] + \\
&\quad \left[-\hat{f}_t'(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{F}^{0\prime}\varepsilon_i^0 \right] + \varepsilon_{it} \\
&= R_{1,it} + R_{2,it} + R_{3,it} + \varepsilon_{it}, \qquad t = 1, 2, \cdots, T; \ \forall i \in \mathcal{T}.
\end{aligned}$$

$\mathbb{E}_\varepsilon(\varepsilon_{it}|D, X, \Lambda, F) = 0$ by Assumption 2. Following a similar logic in Abadie, Diamond and Hainmueller (2010), $R_{3,it}$ can be written as:

$$R_{3,it} = -\sum_{s=1}^{T_0} \hat{f}_t' \left(\sum_{l=1}^{T_0} \hat{f}_l \hat{f}_l' \right)^{-1} \hat{f}_s \varepsilon_{is}$$

in which $\left(\sum_{l=1}^{T_0} \hat{f}_l \hat{f}_l' \right)^{-1}$ is symmetric and positive definite. Applying the Cauchy-Schwarz Inequality, we have $|\hat{f}_t' \left(\sum_{l=1}^{T_0} \hat{f}_l \hat{f}_l' \right)^{-1} \hat{f}_s| \leq O(T_0^{-1})$. Because the second moment for ε_{it} exists (Assumption 3), applying the Rosenthal's Inequality, we have:

$$\mathbb{E}_\varepsilon(|R_{3,it}|^2|D, X, \Lambda, F) \leq O(T_0^{-2}) \sum_{s=1}^{T_0} \mathbb{E}|\varepsilon_{is}|^2 = O(T_0^{-1}).$$

Hence, $\mathbb{E}_\varepsilon(|R_{3,it}| \, |D, X, \Lambda, F) \leq O(T_0^{-1/2})$, which means the bias from $R_{3,it}$ is bounded by a function that goes to zero as the number of pre-treatment periods grows.

Next, we investigate biases from $R_{1,it}$ and $R_{2,it}$. $R_{1,it}$ is the source of bias from imprecise estimation of β, which results in both a direct effect on the amount of bias through x_{it} and an indirect effect through the estimation of the factor loading λ_i. $R_{2,it}$ is the source of bias directly from the influence of the factors $\lambda_i' f_t$. Our objective is to characterize (and bound) both $\mathbb{E}_\varepsilon(R_{1,it})$ and $\mathbb{E}_\varepsilon(R_{2,it})$. By Lemma 2 and $\|x_{it}' - \hat{f}_t'(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{F}^{0\prime}X_i^0\| = O_p(1)$ and

$$|R_{1,it}| = O_p(\|\beta - \hat{\beta}\|) = O_p(N_{co}^{-1}) + O_p(T^{-1}) + o_p\left((N_{co}T)^{-1/2}\right).$$

Using both Lemma 3 and Lemma 4, we have:

$$\begin{aligned}
\|f_t - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{f}_t\| &\leq \|f_t - F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t\| + \|F'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{f}_t - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{f}_t\| \\
&= O_p(N_{co}^{-1/2}) + O_p(T_0^{-1/2}).
\end{aligned}$$

Therefore, $|R_{2,it}| \leq \|\lambda_i\| \|f_t - F^{0\prime}\hat{F}^0(\hat{F}^{0\prime}\hat{F}^0)^{-1}\hat{f}_t\| = O_p(N_{co}^{-1/2}) + O_p(T_0^{-1/2})$.

As both N_{co} and T_0 go to infinity, $O_p(N_{co}^{-1/2}) + O_p(T_0^{-1/2})$ becomes $o_p(1)$. In other words, $R_{1,it} + R_{2,it}$ is bounded in probability by a function that goes to zero. By the moment conditions specified in Assumptions 3 and 4, $R_{1,it} + R_{2,it}$ is uniformly integrable, therefore, convergence in probability implies convergence in means (DasGupta 2008, Ch. 6), i.e., $\mathbb{E}_\varepsilon(|R_{1,it} + R_{2,it}|\ |D, X, \Lambda, F) = o(1)$, as $N_{co}, T_0 \to \infty$. Therefore,

$$\mathbb{E}_\varepsilon(\widehat{ATT}_t - ATT_t | D, X, \Lambda, F) = N_{tr}^{-1} \cdot o(1),$$

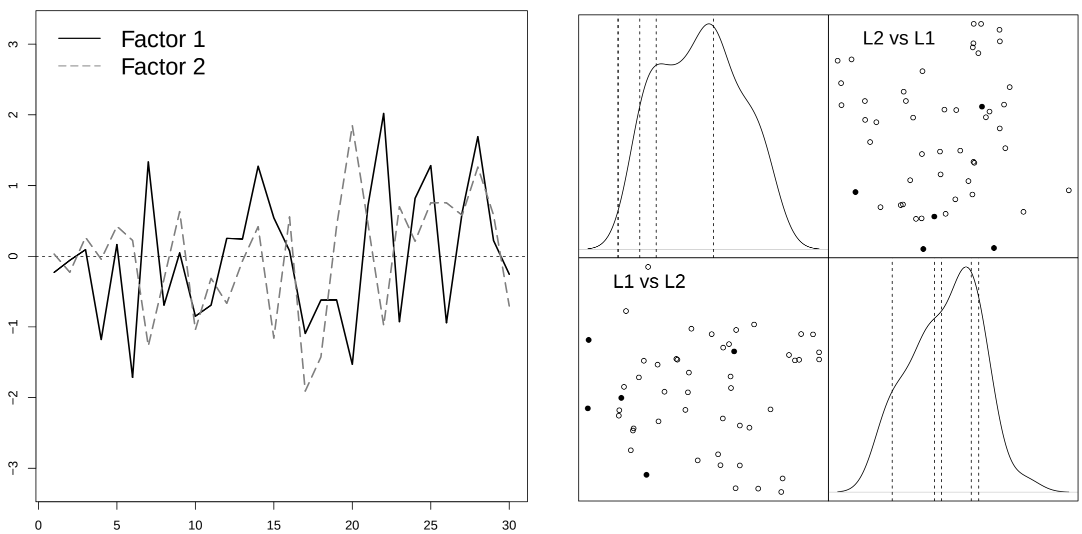
In other words, the bias of the estimator shrinks to zero as both N_{co} and T_0 increase.

B Simulation Results

B.1 Additional Results on The Simulated Example

Figures A2 shows the estimated factors and factor loadings based on the simulated sample. In Figure A2(b), plots on the diagonal show the distributions of estimated factor loadings for the control group (density plot) and the treatment group (dashed lines); plots off the diagonal are scatter-plots of the factor loadings in which solid and hollow circles represent units in the treatment group and the control group, respectively.

Figure A2. Estimated Factors and Factor Loadings: A Simulated Sample

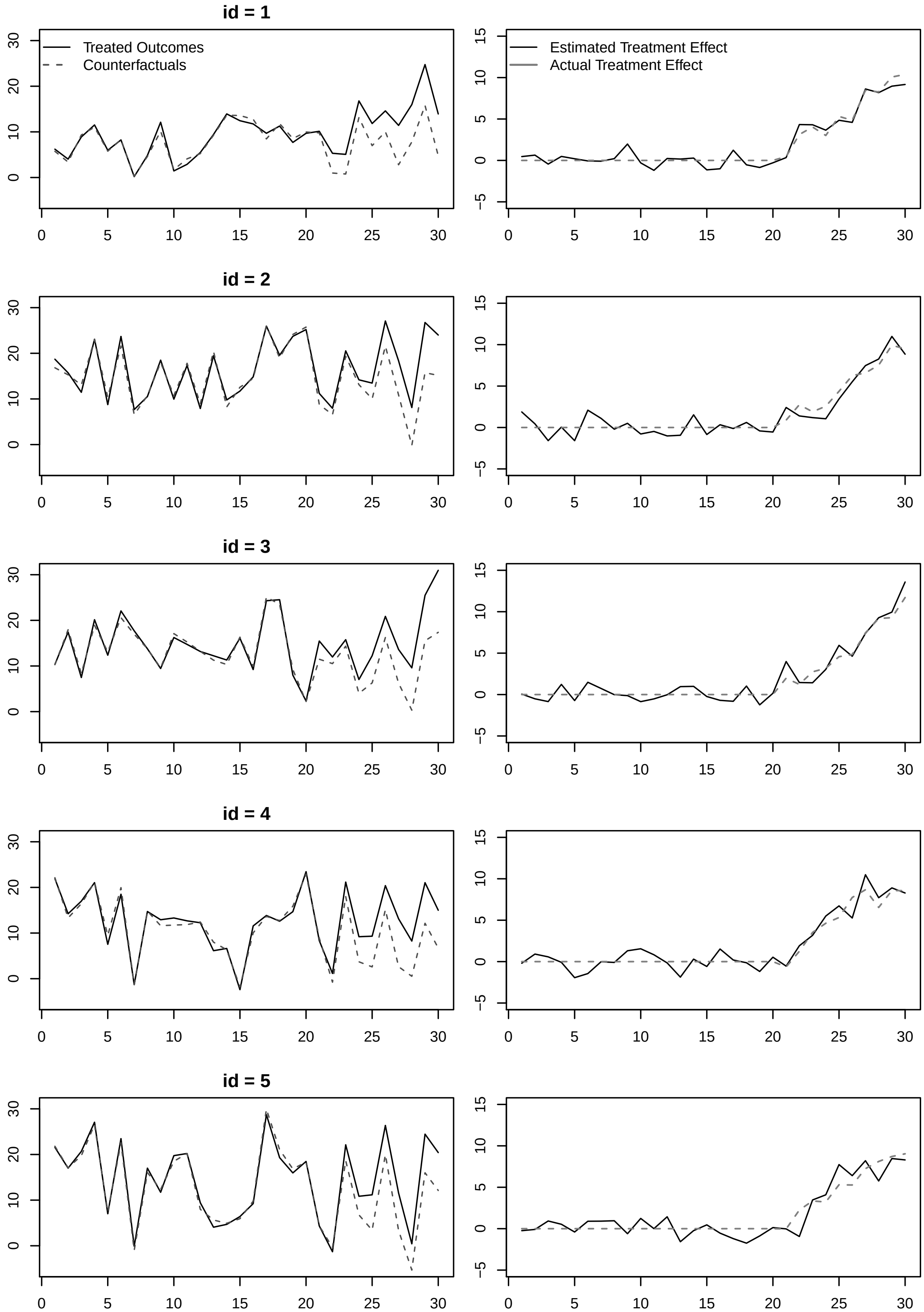


(a) Estimated factors (b) Estimated factor loadings

Figure A3 presents the imputed counterfactual and individual treatment effect for each treated units.

Figure A3. Estimated Individual Treatment Effect: A Simulated Sample

$N_{tr} = 5, N_{co} = 45, T = 30, T_0 = 10$



B.2 Finite Sample Properties

Table A1 explores the finite sample properties of the GSC estimate. It presents the bias, standard deviation, and root mean square error of $\widehat{ATT}_{T_0+5}$ with different combinations of N_{tr}, N_{co}, and T_0. It shows that the GSC estimator has limited bias when both N_{co} and T_0 are large. Each set of numbers is based on 5,000 simulated samples.

The data generating process (DGP) is the same as described in the main text, with $w = 0.8$ which produces positive correlations among the treatment indicator, factor loadings, the regressors. For each set of simulations, the treatment effects, regressors, factors, and factor loadings are drawn only once, while the error terms are drawn repeatedly.

Table A1. Finite Sample Properties

		$N_{tr} = 1$			$N_{tr} = 5$			$N_{tr} = 20$		
T_0	N_{co}	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
15	40	0.023	1.163	1.163	0.053	0.589	0.591	0.013	0.375	0.375
15	80	-0.031	1.159	1.159	0.017	0.535	0.536	0.010	0.310	0.310
15	120	0.008	1.143	1.143	0.010	0.524	0.524	0.003	0.294	0.294
15	200	-0.004	1.154	1.154	0.011	0.518	0.518	0.004	0.278	0.278
30	40	-0.007	1.089	1.089	0.046	0.538	0.540	0.013	0.351	0.351
30	80	0.012	1.074	1.074	0.021	0.504	0.505	0.011	0.293	0.294
30	120	-0.000	1.072	1.071	0.024	0.494	0.495	0.012	0.275	0.275
30	200	0.005	1.083	1.083	0.008	0.487	0.487	0.002	0.263	0.263
50	40	-0.014	1.072	1.072	0.031	0.519	0.520	0.004	0.342	0.342
50	80	0.014	1.039	1.039	0.016	0.497	0.498	0.006	0.277	0.277
50	120	-0.014	1.039	1.039	0.003	0.475	0.475	0.005	0.261	0.261
50	200	0.013	1.032	1.032	0.016	0.468	0.469	0.005	0.254	0.254

B.3 Comparison with the DID Estimator

Table A2 compares the GSC and DID estimates (i.e., two-way fixed effects plus 10 dummies indicating 10 periods after T_0. It presents the bias, standard deviation, and root mean square error of $\widehat{ATT}_{T_0+5}$ with different combinations of N_{co} and w. It shows that, in the presence time-varying confounders, even when the treatment is randomly assigned ($w = 1$), the DID estimator can have considerable bias due to the imbalance of λ_i between treatment and control groups—the size of the bias depends on the particular draw of λ_i (and hence X_i). Even when the treatment is not randomly assigned ($w < 1$), the bias of GSC estimates remains small.

The data generating process (DGP) is the same as described in the main text. For each set of simulations, the treatment effects, regressors, factors, and factor loadings are drawn only once, while the error terms are drawn repeatedly. In the case of $w = 1$, the bias of the DID estimator will go away once we marginalize out λ_i (i.e., if we draw λ_i for each unit repeatedly).

Table A2. Comparision with the Difference-in-Differences Esitmator

				GSC			DID		
T_0	N_{tr}	N_{co}	w	Bias	SD	RMSE	Bias	SD	RMSE
15	20	40	1.00	0.023	0.326	0.327	0.462	0.282	0.541
15	20	80	1.00	0.014	0.287	0.287	0.068	0.256	0.264
15	20	120	1.00	0.007	0.282	0.283	0.363	0.253	0.443
15	20	200	1.00	0.009	0.271	0.271	0.409	0.243	0.476
15	20	40	0.80	0.024	0.378	0.378	-0.012	0.285	0.285
15	20	80	0.80	0.004	0.310	0.310	0.037	0.258	0.261
15	20	120	0.80	0.007	0.293	0.293	0.240	0.248	0.345
15	20	200	0.80	0.013	0.280	0.281	0.281	0.244	0.372
15	20	40	0.60	0.000	0.494	0.494	-0.031	0.286	0.287
15	20	80	0.60	0.003	0.368	0.368	-0.257	0.257	0.364
15	20	120	0.60	0.008	0.354	0.354	-0.198	0.253	0.321
15	20	200	0.60	0.011	0.317	0.317	-0.237	0.244	0.340

B.4 Comparison with the IFE Estimator

The GSC estimator corrects biases of the IFE estimator when the treatment effect is heterogeneous. When the treatment effect is constant, the IFE estimator is more efficient because it uses information of both the treatment and control groups to estimate covariate coefficients and factors while the proposed method uses the control group information only. When the treatment effect is heterogeneous, however, using IFE models that assume constant treatment effect will lead to biased estimates because heterogeneities of the treatment effect will cause inconsistent estimation of the factor space.

Results from Monte Carlo exercises are consistent with the above intuition. Table A3 compares the performances of the GSC estimator and the IFE estimator. It presents the bias, standard deviation, and root mean square error of $\widehat{ATT}_{T_0+5}$ with different combinations of N_{co} and the standard deviation of the treatment effect $SE(\delta_{it})$. The DGP is as specified in the main text with w set to 1. The IFE model to be estimated uses the following specification:

$$Y_{it} = \sum_{t=T_0+1}^{T} \delta_t D_{it} + x'_{it}\beta + \lambda'_i F_t + \alpha_i + \xi_t + \varepsilon_{it},$$

which allows the treatment effects to be different over time. Table A3 shows that (1) when the treatment effect is constant across units, both estimators have limited bias. The efficiency gain of the IFE estimator is small when N_{tr} is relatively small. (2) when the treatment effect is heterogeneous, the bias of the GSC estimates remains small while the bias of of the IFE estimates is no longer negligible.

Table A3. Comparision with the Interactive Fixed-Effect Estimator

				GSC			IFE		
T_0	N_{co}	N_{tr}	$SE(\delta_{it})$	Bias	SD	RMSE	Bias	SD	RMSE
15	20	40	0	0.027	0.329	0.330	0.000	0.328	0.328
15	20	80	0	0.007	0.291	0.291	-0.004	0.290	0.290
15	20	120	0	0.003	0.285	0.285	-0.003	0.285	0.285
15	20	200	0	0.008	0.268	0.268	0.003	0.268	0.268
15	20	40	5	0.019	0.330	0.331	0.136	1.440	1.446
15	20	80	5	0.015	0.287	0.287	-0.547	0.469	0.720
15	20	120	5	0.008	0.275	0.275	0.107	0.269	0.290
15	20	200	5	0.007	0.271	0.271	-0.120	0.282	0.306

B.5 Comparison with the Synthetic Matching Method

The synthetic control method proposed by Abadie, Diamond and Hainmueller (2010) requires that both covariates and factor loadings of the treated unit is in the convex hull of those of the donors from the control group. The method may fail to construct a synthetic control unit when this requirement is not met. In this way, it safeguards against unwarranted extrapolations that may lead to biased estimates of the treatment effect.

The GSC method, however, does not have this requirement–in this sense, it is less conservative in terms of imputing treated counterfactuals. First, like DID, it allows for an intercept shift when additive unit fixed effects are assumed. Second, it incorporates observable covariates by imposing parametric assumptions. Third, in the lack of common support of factor loadings between the treated and control groups, it *extrapolates* the influence of the factors on the treated outcome based on the assumed model. When the model is correct, the GSC estimator is expected to be more efficient than the synthetic matching method because it potentially uses more information: (1) no control units are discarded and even negative correlations between the treated and control units are used for the prediction of treated counterfactuals; (2) when the model specifies more than one unobserved factors, a control unit at different time periods is assigned different weights. To be more precise, the control units are first decomposed into several components (factors) and these components are re-weighted to produce treated counterfactuals. When the model is incorrect, however, such extrapolations cause biases. Therefore, when applying the GSC method, it is helpful to plot the estimated factors and factor loadings to avoid excessive extrapolations.

Table A4 compares the GSC method with the original synthetic control method (labelled as `Synth`) and confirms our expectation. It presents the bias, standard deviation, and root mean square error of $\widehat{ATT}_{T_0+5}$ with different combinations of r and w, a parameter that characterizes the overlap of support of factor loadings (including unit fixed effects) between the treated and control units. The DGP is the same as described in the main text. Note that for each set of simulations, the factors are drawn only once, while the treatment effect, regressors, factor loadings, and error terms are drawn repeatedly.

Table A4 shows that when the model is correct and the treated and control units shares a common support of factor loadings, both methods have limited bias while the GSC estimator is more efficient. In our setup, there is a small chance that the synthetic matching method may fail to construct a synthetic control unit. The chances may become high when N_r is bigger. As the overlap of support between treated and control units diminishes, significant bias shows up for the original synthetic matching method while the bias of the GSC method remains small.

Table A4. Comparison with the Synthetic Control Estimator (ADH 2010)

T_0	N_{tr}	N_{co}	r	w	GSC				`Synth`			
					Bias	SD	RMSE	Fail	Bias	SD*	RMSE*	Fail
15	1	40	1	1.00	-0.010	1.494	1.107	0.000	-0.011	1.739	1.417	0.013
15	1	40	2	1.00	0.000	1.571	1.190	0.000	-0.022	2.029	1.738	0.008
15	1	40	3	1.00	-0.003	1.581	1.220	0.000	0.021	2.368	2.151	0.014
15	1	40	4	1.00	0.013	1.610	1.253	0.000	-0.013	2.345	2.122	0.016
15	1	40	2	0.75	0.014	1.595	1.234	0.000	0.707	2.096	1.975	0.016
15	1	40	2	0.50	0.026	1.602	1.257	0.000	1.331	2.327	2.489	0.014
15	1	40	2	0.25	0.000	1.729	1.391	0.000	1.630	2.492	2.803	0.016
15	1	40	2	0.00	0.033	1.822	1.521	0.000	2.127	2.610	3.199	0.012

B.6 Model Selection

So far we have shown that when the factor model is correctly specified, the GSC estimator performs well in small samples and have advantages over the DID estimator, the synthetic matching method, and the IFE estimator under various circumstances. In this section, we investigate whether the cross-validation scheme proposed earlier in this paper is able to select the correct number of factors when it is unknown and whether the GSC method is robust to a misspecified number of factors.

Table A5. Choices of the Number of Factors

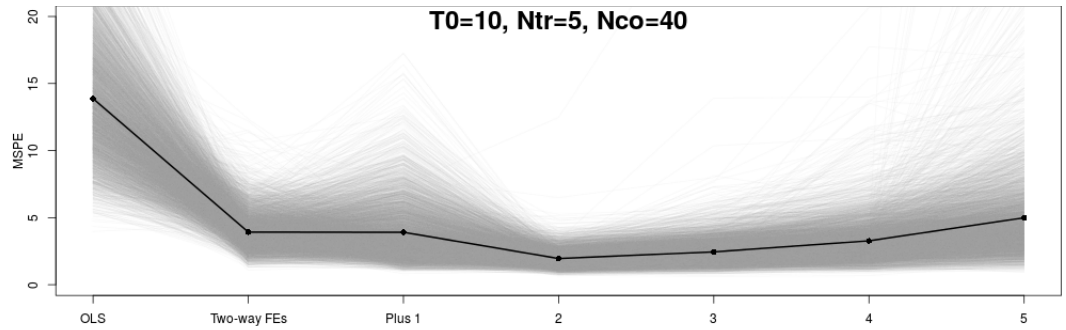
		$N_{tr} = 5$	$N_{tr} = 20$	$N_{tr} = 40$
T_0	N_{co}	r✓	r✓	r✓
10	40	0.801	0.938	0.953
30	40	0.921	0.985	0.991
50	40	0.943	0.990	0.996
15	40	0.879	0.976	0.991
15	80	0.896	0.992	0.998
15	120	0.895	0.995	0.999

Note: Each number is based on 5,000 simulated samples.

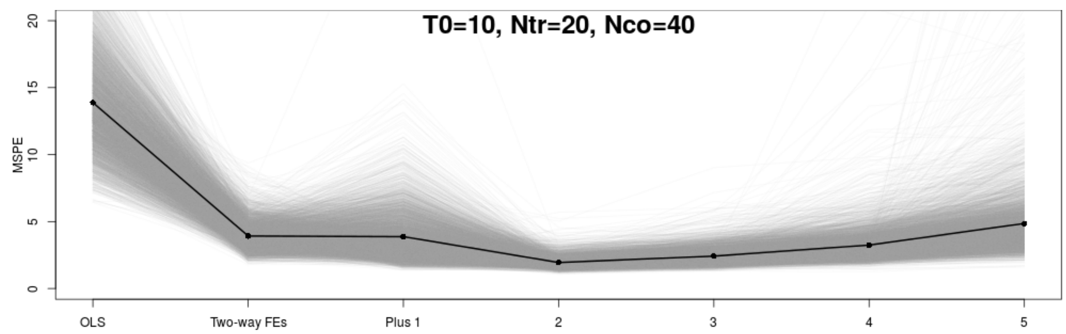
We conduct simulations using the same DGP specified in Equation (3) (with $w = 0.5$) and let the algorithm choose the number of factors automatically. Table A5 shows the percentage of correct choices of the number of factors with different sample size from 5,000 simulations for each case. It suggests that when the sample is reasonably large, with a high chance the cross-validation scheme can choose the number of factors correctly. For example, when $T_0 = 30$, $N_{co} = 40$ and $N_{tr} = 5$, the cross-validation algorithm correctly chooses the number of factors 92.1% of the time; the number increases to 98.5% when $N_{tr} = 20$. Note that the number of treated units N_{tr} matters because a larger treatment group provide more data for validation.

In Figure A4, we choose four cases (with different combinations of T_0, N_{co}, and N_{tr}) and plot the MSPEs of six models, including pooled OLS, the two-way fixed effects model, and the GSC method with 1 to 5 factors (shown on the x-axis), with 5,000 simulations for each case. Results from all simulations are represented with gray lines. The black solid line shows the median MSPE of 5,000 simulations with each model. These figures show that the median MSPE is always the lowest with the correct number of factors, i.e., $r = 2$.

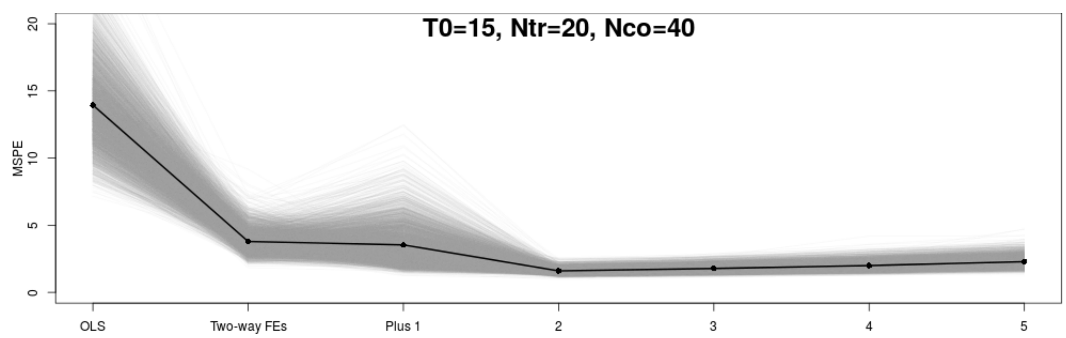
Figure A4. Choice of the Number of Factors: Four Cases



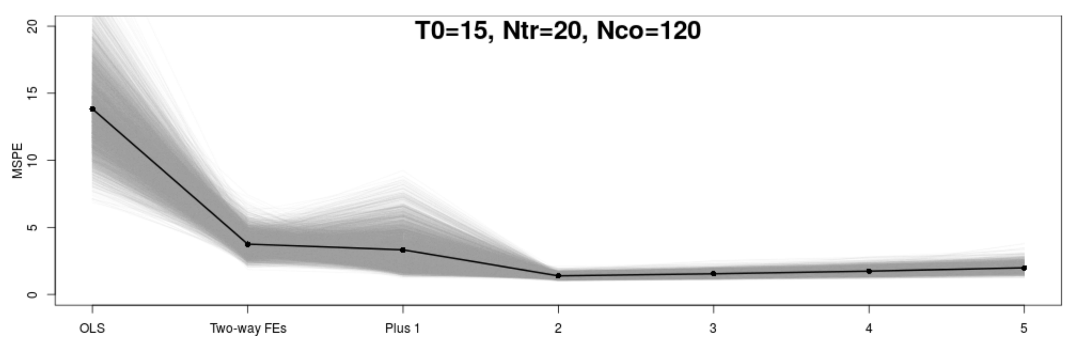
(a)



(b)



(c)



(d)

C Additional Results on EDR Laws and Turnout

Table A6 lists the years during which EDR laws were enacted and first took effect in presidential elections for the 9 treated states.

Table A6. State EDR Laws

State	Enacted	Took effect
Maine	1973	1976
Minnesota	1974	1976
Wisconsin	1975	1976
Wyoming	1994	1996
Idaho	1994	1996
New Hampshire	1996	1996
Montana	2005	2008
Iowa	2007	2008
Connecticut	2012	2012

Figure A5 shows the raw data of state-level turnout rates (%) in US presidential elections in 47 states from 1920 to 2012. Turnout rates of 38 states that had not adopted EDR laws (controls) are in gray. For the 9 states in which EDR laws took effect before 2012 (treated), the pre- and post-EDR periods are represented by black solid lines and black dashed lines, respectively.

Figure A5. EDR and State-level Turnout: Raw Data

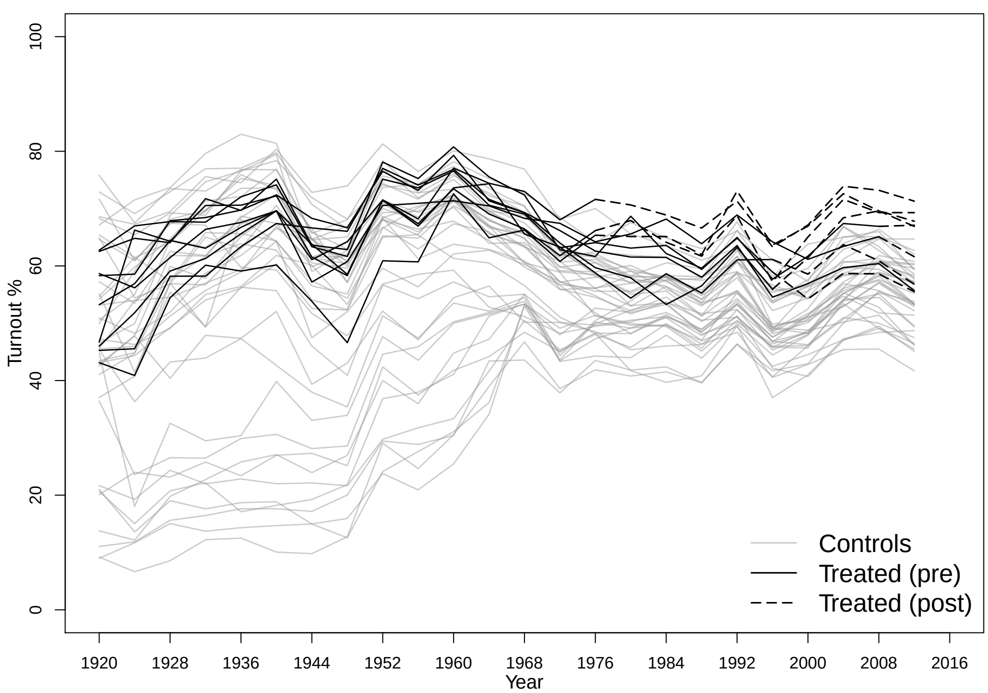
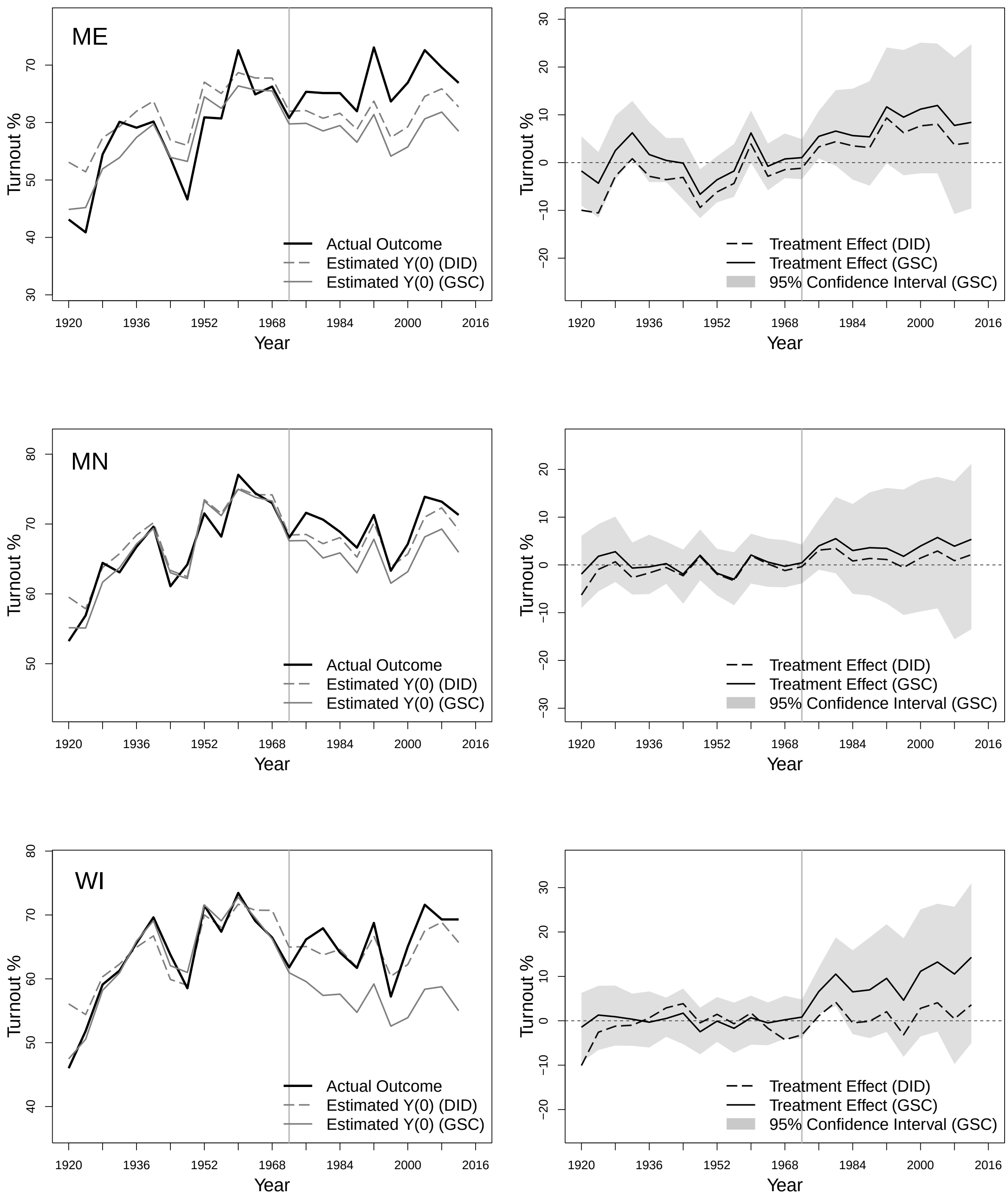
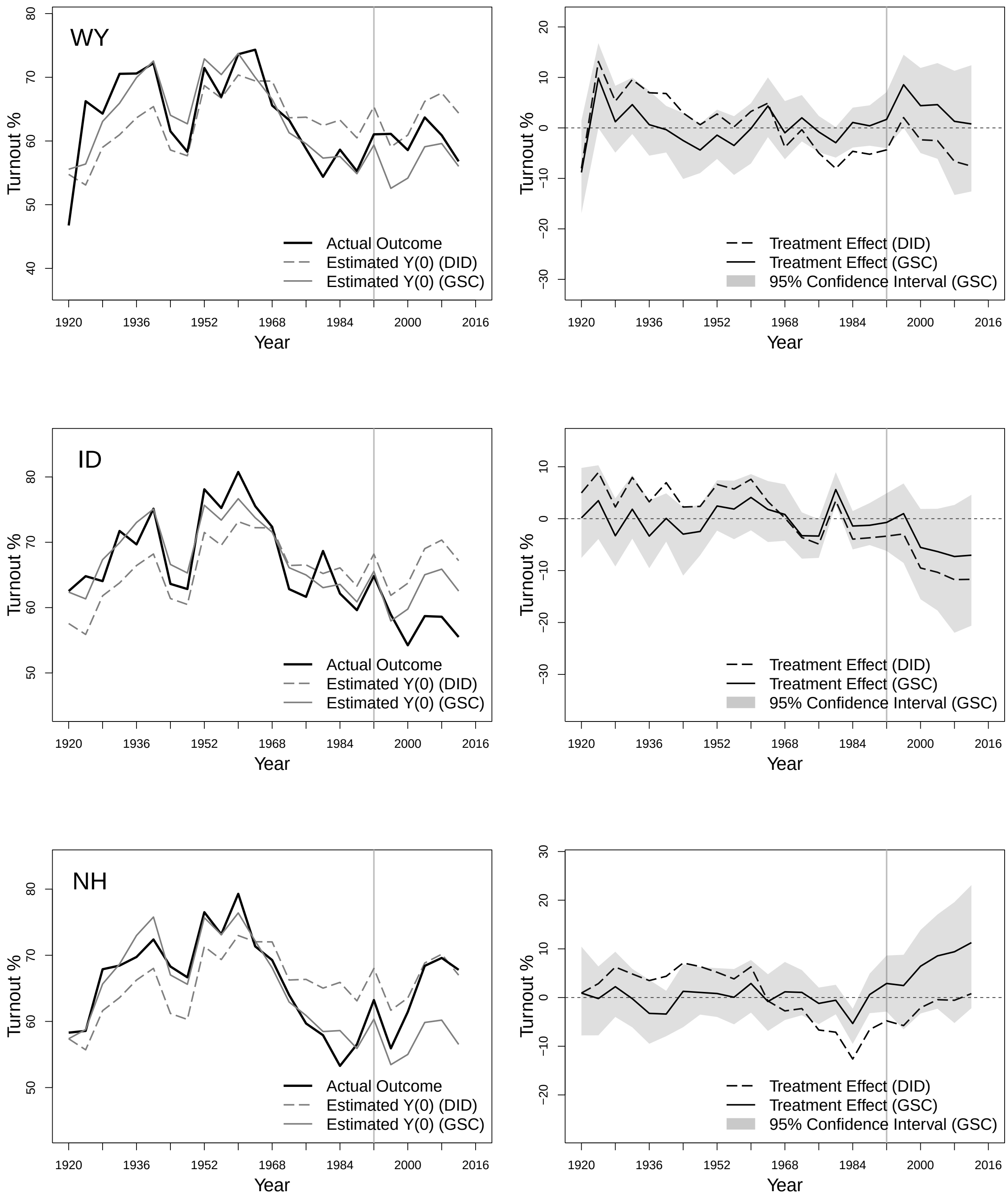
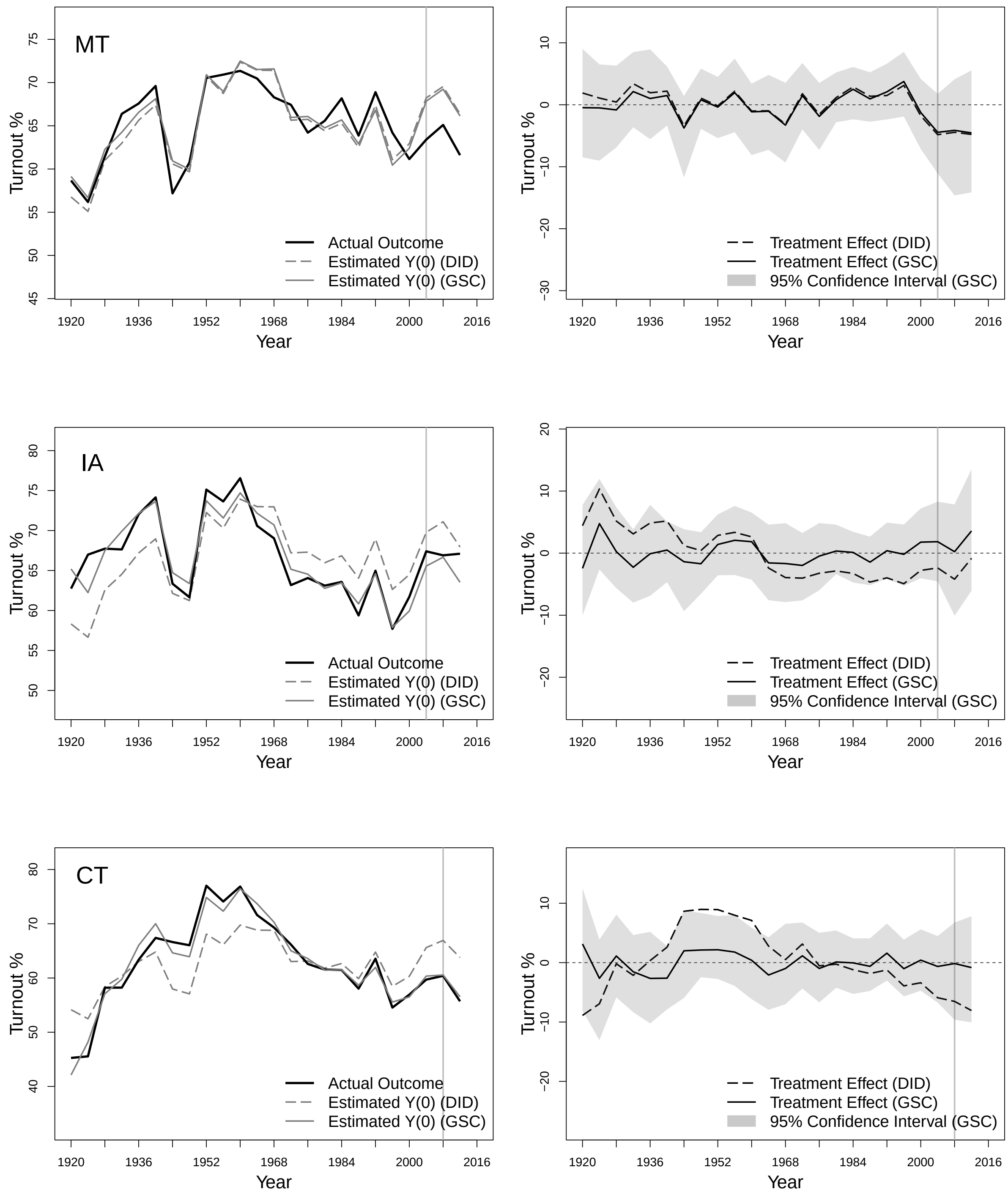


Figure A6 shows the imputed counterfactuals and estimated individual treatment effect produced by both the DID and GSC estimators for each of the 9 states that enacted EDR laws before the 2012 presidential election. The shades represent the 95% confidence intervals for the treatment effects produced by the GSC method.

Figure A6. The Effect of EDR on Turnout: Individual Cases




WY
ID
NH
Turnout %
Year
1920
1936
1952
1968
1984
2000
2016
Actual Outcome
Estimated Y(0) (DID)
Estimated Y(0) (GSC)
Treatment Effect (DID)
Treatment Effect (GSC)
95% Confidence Interval (GSC)


MT
IA
CT
Turnout %
Year
1920
1936
1952
1968
1984
2000
2016
Actual Outcome
Estimated Y(0) (DID)
Estimated Y(0) (GSC)
Treatment Effect (DID)
Treatment Effect (GSC)
95% Confidence Interval (GSC)

References

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490):493–505.

Bai, Jushan. 2003. "Theory for Factor Models of Large Dimensions." *Econometrica* 71(1):135–137.

Bai, Jushan. 2009. "Panel Data Models with Interactive Fixed Effects." *Econometrica* 77:1229–1279.

DasGupta, Anirban. 2008. *Asymptotic Theory of Statistics and Probability*. New York, NY: Springer.