Appendix to "Estimation of Average Treatment Effects Using Panel Data when Treatment Effect Heterogeneity Depends on Unobserved Fixed Effects"

Shosei Sakaguchi*

December 20, 2019

Abstract

Section A provides proofs of Proposition 3.1–3.3. Section B provides additional Monte Carlo simulation results.

A. Proofs of Propositions 3.1–3.3

This section provides the proofs of Propositions 3.1–3.3. In this section, we denote the vector of the true parameters by $\theta_0 = (\tilde{\alpha}_0^1, \tilde{\alpha}_0^0, \beta_0^{1\prime}, \beta_0^{0\prime}, \gamma_0^1, \tau_{1,0}^{ate}, \cdots, \tau_{T,0}^{ate})'$. The proofs use the moment functions defined in Section 3.2 of the main text.

(Proof of Proposition 3.1.)

For the moment function $g_1(W_i, \theta)$, under Assumptions 2.1 and 2.2 and $\theta = \theta_0$, the following holds:

$$
\begin{aligned}
& E\left[g_1(W_i, \theta_0)\right] \\
=& E\left[\sum_{t=1}^{T} D_{it} \cdot \ddot{X}_{it}^1 (\ddot{Y}_{it}^1 - \ddot{X}_{it}^{1\prime}\beta_0^1)\right] \\
=& \sum_{t=1}^{T} E\left[\ddot{X}_{it}^1 \ddot{u}_{it}^1 \mid D_{it} = 1\right] \cdot P(D_{it} = 1) \\
=& \sum_{t=1}^{T} E\left[\ddot{X}_{it}^1 \left(u_{it}^1 - \frac{\sum_{s=1}^{T} D_{is} u_{is}^1}{\sum_{s=1}^{T} D_{is}}\right) \mid D_{it} = 1\right] \cdot P(D_{it} = 1) \\
=& \sum_{t=1}^{T} E\left[\ddot{X}_{it}^1 E\left[u_{it}^1 \mid D_{i1}, \ldots, D_{iT}, X_{i1}, \ldots, X_{iT}, C_i\right] \mid D_{it} = 1\right] \cdot P(D_{it} = 1)
\end{aligned}
$$

*Department of Economics, University College London, Gower Street, London WC1E 6BT, United Kingdom. Email: s.sakaguchi@ucl.ac.uk.

$$- \sum_{t=1}^{T} E\left[\ddot{X}_{it}^{1}\left(\frac{\sum_{s=1}^{T} D_{is} E\left[u_{is}^{1} \mid D_{i1},\ldots,D_{iT},X_{i1},\ldots,X_{iT},C_i\right]}{\sum_{s=1}^{T} D_{is}}\right) \mid D_{it}=1\right] \cdot P(D_{it}=1)$$
$$=0.$$

The last equality holds by Assumption 2.2 because, under this assumption and for $j = 0, 1$ and $t = 1, \ldots, T$,

$$\begin{aligned}
&E[u_{it}^{j} \mid D_{i1},\ldots,D_{iT},X_{i1},\ldots,X_{iT},C_i] \\
=&E[Y_{it}(j) - E\left[Y_{it}(j) \mid X_{it},C_i\right] \mid D_{i1},\ldots,D_{iT},X_{i1},\ldots,X_{iT},C_i] \\
=&E\left[Y_{it}(j) \mid X_{it},C_i\right] - E\left[Y_{it}(j) \mid X_{it},C_i\right] \\
=&0.
\end{aligned}$$

Hence, under these assumptions, we have

$$E\left[g_1(W_i,\theta_0)\right] = E\left[\sum_{t=1}^{T} D_{it}\cdot \ddot{X}_{it}^{1}\ddot{Y}_{it}^{1}\right] - E\left[\sum_{t=1}^{T} D_{it}\cdot \ddot{X}_{it}^{1}\ddot{X}_{it}1\prime\right]\beta_0^1 = 0.$$

Then, under Assumption 3.1, the value of β^1 that satisfies $E\left[g_1(W_i,\theta_0)\right] = 0$ is uniquely obtained as

$$\beta_0^1 = E\left[\sum_{t=1}^{T} D_{it}\cdot \ddot{X}_{it}^{1}\ddot{X}_{it}^{1\prime}\right]^{-1} E\left[\sum_{t=1}^{T} D_{it}\cdot \ddot{X}_{it}^{1}\ddot{Y}_{it}^{1\prime}\right].$$

In a similar way, it can be shown that, under the same assumptions, the value of β^0 that satisfies $E\left[g_2(W_i,\theta)\right] = 0$ is uniquely obtained as

$$\beta_0^0 = E\left[\sum_{t=1}^{T}(1-D_{it})\cdot \ddot{X}_{it}^{0}\ddot{X}_{it}^{0\prime}\right]^{-1} E\left[\sum_{t=1}^{T}(1-D_{it})\cdot \ddot{X}_{it}^{0}\ddot{Y}_{it}^{0\prime}\right].$$

□

(Proof of Proposition 3.2.)

For the moment function $g_3(W_i,\theta)$, under Assumptions 2.1, 2.2, 2.3, and 2.4 (i) and $\theta = \theta_0$, the following holds:

$$\begin{aligned}
&E\left[g_3(W_i,\theta_0)\right] \\
=&E\left[\sum_{t=1}^{T} 1\left\{D_{it}=1, \sum_{t=1}^{T} D_{it} \neq T\right\}\cdot \left(1, Z_{it}'\right)' \left(Y_{it} - \tilde{\alpha}_0^1 - X_{it}'\beta_0^1 - \gamma_0^1\eta_i^0(\beta_0^0)\right)\right]
\end{aligned}$$

$$
\begin{aligned}
=&\sum_{t=1}^{T} E\left[\left(1, Z_{it}'\right)'\left(Y_{it}(1)-\tilde{\alpha}_0^1-X_{it}'\beta_0^1-\gamma_0^1\eta_i^0(\beta_0^0)\right) \mid D_{it}=1, \sum_{t=1}^{T} D_{it}\neq T\right]\cdot P(D_{it}=1,\sum_{t=1}^{T} D_{it}\neq T)\\
=&\sum_{t=1}^{T} E\left[\left(1, Z_{it}'\right)'\left(u_{it}^1-\gamma_0^1\bar{u}_i^0\right) \mid D_{it}=1, \sum_{t=1}^{T} D_{it}\neq T\right]\cdot P(D_{it}=1,\sum_{t=1}^{T} D_{it}\neq T)\\
=&\sum_{t=1}^{T} E\left[\left(1, Z_{it}'\right)' u_{it}^1 \mid D_{it}=1, \sum_{t=1}^{T} D_{it}\neq T\right]\cdot P(D_{it}=1,\sum_{t=1}^{T} D_{it}\neq T)\\
&-\gamma_0^1\cdot\sum_{t=1}^{T}\sum_{s=1}^{T} E\left[\frac{(1-D_{is})}{\sum_{s=1}^{T}(1-D_{is})}E\left[\left(1, Z_{it}'\right)' u_{is}^0 \mid D_{it}=1, D_{is}=0, \sum_{t=1}^{T} D_{it}\neq T\right]\right.\\
&\left.\mid D_{it}=1, \sum_{t=1}^{T} D_{it}\neq T\right]\cdot P(D_{it}=1,\sum_{t=1}^{T} D_{it}\neq T)\\
=&0.
\end{aligned}
$$

The existence of the subpopulation $\{i : D_{it} = 1, \sum_{t=1}^{T} D_{it} \neq T\}$ is guaranteed by Assumption 2.3. The last equality holds by Assumption 2.4 (i).

Given that $\beta^1 = \beta_0^1$ and Assumption 2.4 (ii) holds, the values of $(\tilde{\alpha}^1, \gamma^1)$ that satisfy this moment condition are uniquely obtained as $(\tilde{\alpha}_0^1, \gamma_0^1)$. By a similar argument, it can be shown that, under the same assumptions, the values of $(\tilde{\alpha}^0, \gamma^1)$ that satisfy the moment condition $E[g_4(W_i, \theta)] = 0$ are uniquely obtained as $(\tilde{\alpha}_0^0, \gamma_0^1)$.

□

(Proof of Proposition 3.3.)

For the moment function $g_{4+t}(W_i, \theta)$ for $t \in \{1, \cdots, T\}$, under Assumptions 2.1 and 2.2 and $\theta = \theta_0$, the following holds for the expectation of its first term:

$$
\begin{aligned}
&E\left[D_{it}Y_{it}-(1-D_{it})\left(\tilde{\alpha}_0^1+X_{it}'\beta_0^1+\gamma^1\eta_i^0(\beta_0^0)\right)\right]\\
=&E\left[D_{it}Y_{it}(1)-(1-D_{it})(Y_{it}(1)-u_{it}^1+\gamma_0^1\bar{u}_i^0)\right]\\
=&E\left[Y_{it}(1)\right]+E\left[(1-D_{it})(u_{it}^1-\gamma_0^1\bar{u}_i^0)\right]\\
=&E\left[Y_{it}(1)\right].
\end{aligned}
$$

The last equality holds because $E\left[(1-D_{it})\left(u_{it}^0-\gamma_0^1\bar{u}_i^0\right)\right] = 0$ holds under Assumption 2.2. Similarly, under the same assumptions and $\theta = \theta_0$, the following holds for the expectation of the second term of $g_{4+t}(W_i, \theta)$:

$$
\begin{aligned}
&E\left[D_{it}\left\{\tilde{\alpha}_0^0+X_{it}'\beta_0^0+\frac{1}{\gamma_0^1}\eta_i^1(\beta_0^1)\right\}-(1-D_{it})Y_{it}\right]\\
=&E\left[D_{it}\left\{Y_{it}(0)-u_{it}^0+\frac{1}{\gamma_0^1}\bar{u}_i^1\right\}-(1-D_{it})Y_{it}(0)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= E\left[Y_{it}(0)\right] + E\left[D_{it}\left\{-u_{it}^0 + \frac{1}{\gamma_0^1}\bar{u}_i^1\right\}\right] \\
&= E\left[Y_{it}(0)\right].
\end{aligned}
$$

Thus, given $(\tilde{\alpha}^0, \tilde{\alpha}^1, \beta^0, \beta^1, \gamma^1) = (\tilde{\alpha}_0^0, \tilde{\alpha}_0^1, \beta_0^0, \beta_0^1, \gamma_0^1)$ and under Assumptions 2.1 and 2.2, $E\left[g_{4+t}(W_i, \theta)\right] = 0$ uniquely holds with

$$
\begin{aligned}
\tau_{t,0}^{ate} &= E[Y_{it}(1)] - E[Y_{it}(0)] \\
&= E\left[D_{it}\left\{Y_{it} - \left(\tilde{\alpha}^0 + X_{it}'\beta^0 + \frac{1}{\gamma^1}\eta_i^1(\beta^1)\right)\right\} + (1 - D_{it})\left\{\left(\tilde{\alpha}^1 + X_{it}'\beta^1 + \gamma^1\eta_i^0(\beta^0)\right) - Y_{it}\right\}\right].
\end{aligned}
$$

□

B. Additional Monte Carlo Simulations Results

This section provides additional Monte Carlo simulation results. I evaluate the finite sample behavior of the proposed estimator as well as the OLS and FE estimators in three cases: (i) fraction of movers is small; (ii) γ_1 is close to zero; and (iii) T is large.

B.1. Case (i): Fraction of movers is small

In this section, we study the case when the fraction of movers is small. In this case, the moment functions $g_3(W_i, \theta)$ and $g_4(W_i, \theta)$ are likely to use a small sample of movers in estimation. Note that these moment functions are derived from the second step of the identification procedure described in Section 3.1 of the main text. Here we use DGPs that are the same as DGPs 1 and 2 in Section 4 of the main text, except for the pairwise covariance between u_{i1}^D and u_{i1}^D and the one between $X_{i1,k}$ and $X_{i2,k}$ (for $k = 1, 2$). Here I change all these pairwise covariances to $K = 0.8, 0.9$, or 0.98. We refer to the DGP that is same as DGPj ($j = 0, 1$) in Section 4 but the pairwise covariances are equal to $K = 0.8, 0.9$, and 0.98 as DGP$j(1)$, DGP$j(2)$, and DGP$j(3)$, respectively. The ratios of movers in DGP$j(1)$, DGP$j(2)$, and DGP$j(3)$, for each $j = 0, 1$, are about 16.5%, 11.6%, and 5.2%, respectively.

Tables B.1 and B.2 show the results of 1,000 simulations, where the same estimators as in Section 4 of the main text are used. We use sample size $N = 500$ and 800 for all the DGPs.[1] The results show that the SDs and RMSEs of the proposed estimator as well as those of the FE estimator increase as the fraction of movers decreases. But, in the current setting, the performance of the proposed estimator is still compatible with that of the FE estimator even in DGP1(k) for each $k = 1, 2$, and 3.

[1] I also tried the simulation with $N = 200$, but the results could not be obtained due to the singularity of the weighting matrix of the GMM for some finite simulated samples.

B.2. Case (ii): γ^1 is close to zero

We next study the case when γ^1 is close to zero. Recall that γ^1 is the ratio of unobserved fixed effects between the treated and untreated potential outcome equations. In this case, $1/\gamma^1$ in the moment function $g_4(W_i, \theta)$ takes very large value, which might make the proposed estimator vulnerable. We here consider two kinds of DGPs (DGPs 3 and 4), with $T = 2$, which are the same as DGPs 1 and 2 in Section 4 of the main text, except for the value of γ^1. I set γ^1 equal to 0.1 in DGP3 and 0.05 in DGP4. In DGP3, the true values of the ATEs in periods 1 and 2 are 0.3 and 1.3, respectively; in DGP4, the true values of the ATEs in periods 1 and 2 are 0.05 and 1.05, respectively.

Tables B.3 and B.4 show the results of 1,000 simulations, where we use the same estimators and the same sample sizes as in Section 4. The results show that the SDs and RMSEs of the proposed estimator increase as γ^1 becomes close to zero, while those of the OLS and FE estimators do not change so much. In DGP4, the proposed estimator shows higher RMSEs than the OLS and FE estimators. This implies that the proposed estimator is not superior to the OLS and FE estimators when γ^1 is too close to zero. Note also that, in both DGPs, the performance of the OLS estimator is better than that of the FE estimator in terms of RMSEs.

B.3. Case (iii): T is large

Lastly, we study the case when T is large. In this case, the number of observations becomes large and the sample is more likely to have many movers, whereas the number of moments and parameters to be estimated also become large. We consider the cases when $T \in \{10, 20, 30\}$ and $N \in \{200, 500, 800\}$. In the simulation, we use the following two outcome models: For $t = 1, \ldots, T$,

$$\text{(Model 1)} \quad Y_{it} = \begin{cases} 2 + 0.1T + X_{it,1} - 0.5X_{it,2} + C_i + u_{it}^0 & \text{if } D_{it} = 0 \\ 1 + 0.15T + 2X_{it,1} + 0.5X_{it,2} + C_i + u_{it}^1 & \text{otherwise} \end{cases},$$

$$\text{(Model 2)} \quad Y_{it} = \begin{cases} 2 + 0.1T + X_{it,1} - 0.5X_{it,2} + C_i + u_{it}^0 & \text{if } D_{it} = 0 \\ 1 + 0.15T + 2X_{it,1} + 0.5X_{it,2} + 3C_i + u_{it}^1 & \text{otherwise} \end{cases}.$$

Aside from including time trend effects, these two models are the same as the outcome models in DGPs 1 and 2 in Section 4 of the main text, respectively. We use the same form of the treatment assignment model as that used in Section 4: $D_{it} = 1\{-2 - X_{it,1} + X_{it,2} + C_i + u_{it}^D \geq 0\}$. I set the observe covariates and unobserved fixed effects $(X_{it,1}, X_{it,2}, C_i)$ as $X_{it,1} = 0.7 + 0.3C_i + \eta_{it}^1$, $X_{it,2} = 1 + \eta_{it}^2$, and $C_i \sim i.i.d. N(1, 1)$, where $\eta_{it}^j = 0.3\eta_{i,t-1}^j + \epsilon_{it}^j$, $\epsilon_{it}^j \sim i.i.d.\ N(0, 1)$, and $\eta_{i,0}^j \sim i.i.d.\ N(0, 1/0.91)$ for $j = 0, 1$. For the disturbance variables $(u_{it}^0, u_{it}^1, u_{it}^D)$, I set that $(u_{it}^0, u_{it}^1)' \sim 0.3 \cdot \left(u_{i,t-1}^0, u_{i,t-1}^1\right)' + \left(\nu_{it}^0, \nu_{i,t}^1\right)'$ and $u_{it}^D = 0.3u_{i,t-1}^D + \nu_{it}^D$, where

$$\left(\nu_{it}^0, \nu_{i,t}^1\right)' \sim i.i.d.\ N\left(\mathbf{0}, \begin{pmatrix} 1 & 0.5 \\ & 1 \end{pmatrix}\right),\ \left(u_{i,0}^0, u_{i,0}^1\right)' \sim i.i.d.\ N\left(\mathbf{0}, \frac{1}{0.91}\begin{pmatrix} 1 & 0.5 \\ & 1 \end{pmatrix}\right),$$

and $u_{i0}^D \sim i.i.d.\ N(0, 1/0.91)$. Further, I set an exogenous variable W_{it} as $W_{it} = C_i + u_{it}^W$ where $u_{it}^W = 0.3u_{i,t-1}^W + \epsilon_{it}^W$, $u_{i,0}^W \sim i.i.d.\ N(0, 1/0.91)$, and $\epsilon_{it}^W \sim i.i.d.\ N(0, 1)$. Under Models 1 and 2, the true values of the ATEs in period t are $1 + 0.05t$ and $3 + 0.05t$, respectively.

Tables B.5 and B.6 show the results of 1,000 simulations under Models 1 and 2, respectively. The results show that the performance of the proposed estimator is likely to improve as T increases, although there seems to be no clear improvement under Model 1 when comparing the cases $T = 10$ and $T = 20$. This implies that the advantage of large T surpasses the disadvantage of that in the current setting. Note, however, that the results do not guarantee that this outperformance still holds when T takes a much larger number. The results for the FE estimator do not show clear improvement under both models.

Table B.1: Monte Carlo Simulation Results: DGP1(1)–DGP1(3)

		DGP1(1)			DGP1(2)			DGP1(3)		
		Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
N=500	ATE for 1st period									
	OLS	1.164	0.161	1.175	1.165	0.163	1.177	1.168	0.160	1.179
	FE	-0.007	0.195	0.195	-0.011	0.217	0.217	-0.001	0.313	0.313
	Proposed Estimator									
	(1)	0.025	0.156	0.158	0.022	0.173	0.175	0.027	0.211	0.213
	(2)	0.021	0.148	0.149	0.022	0.162	0.163	0.031	0.194	0.196
	(3)	0.032	0.137	0.140	0.028	0.148	0.151	0.040	0.176	0.180
	(4)	0.027	0.139	0.141	0.025	0.155	0.157	0.029	0.187	0.189
	(5)	0.029	0.157	0.159	0.026	0.165	0.166	0.026	0.203	0.204
	ATE for 2nd period									
	OLS	1.169	0.166	1.180	1.171	0.167	1.183	1.163	0.165	1.174
	FE	-0.002	0.194	0.194	-0.005	0.216	0.216	-0.004	0.315	0.315
	Proposed Estimator									
	(1)	0.040	0.164	0.168	0.037	0.177	0.180	0.032	0.228	0.230
	(2)	0.029	0.152	0.154	0.029	0.167	0.169	0.028	0.206	0.208
	(3)	0.040	0.147	0.153	0.040	0.155	0.160	0.039	0.196	0.200
	(4)	0.030	0.147	0.150	0.032	0.163	0.166	0.027	0.206	0.208
	(5)	0.035	0.159	0.162	0.032	0.171	0.173	0.026	0.221	0.222
N=800	ATE for 1st period									
	OLS	1.156	0.267	1.186	1.162	0.127	1.168	1.166	0.133	1.174
	FE	-0.003	0.153	0.153	-0.002	0.168	0.168	0.001	0.246	0.246
	Proposed Estimator									
	(1)	0.022	0.128	0.130	0.021	0.142	0.144	0.026	0.175	0.177
	(2)	0.025	0.116	0.119	0.027	0.127	0.129	0.020	0.158	0.159
	(3)	0.028	0.108	0.112	0.027	0.120	0.123	0.024	0.152	0.153
	(4)	0.026	0.113	0.116	0.027	0.125	0.127	0.024	0.157	0.159
	(5)	0.023	0.122	0.124	0.029	0.132	0.135	0.029	0.171	0.174
	ATE for 2nd period									
	OLS	1.156	0.126	1.163	1.157	0.128	1.164	1.165	0.133	1.173
	FE	-0.004	0.154	0.154	-0.003	0.172	0.172	0.002	0.245	0.245
	Proposed Estimator									
	(1)	0.035	0.128	0.133	0.032	0.147	0.150	0.033	0.185	0.188
	(2)	0.030	0.121	0.124	0.032	0.134	0.137	0.027	0.166	0.168
	(3)	0.038	0.118	0.124	0.037	0.133	0.138	0.030	0.164	0.167
	(4)	0.032	0.121	0.125	0.031	0.137	0.141	0.029	0.171	0.173
	(5)	0.029	0.127	0.130	0.036	0.142	0.146	0.034	0.184	0.187

Note: Bias, SD, and RMSE are the mean bias, standard deviation, and the root mean squared error of the estimates across the simulations, respectively. For the Proposed Estimator, the rows (1), (2), (3), (4), and (5) report the results of the proposed estimator when Z_{it} is $\{W_{it}\}$, $\{X_{it,1}\}$, $\{X_{it,2}\}$, $\{X_{it,1}, X_{it,2}\}$, and $\{W_{it}, X_{it,1}, X_{it,2}\}$, respectively.

Table B.2: Monte Carlo Simulation Results: DGP2(1)–DGP2(3)

		DGP2(1)			DGP2(2)			DGP2(3)		
		Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
N=500	ATE for 1st period									
	OLS	2.875	0.303	2.891	2.887	0.298	2.902	2.858	0.308	2.875
	FE	0.867	0.280	0.911	0.878	0.309	0.930	0.865	0.452	0.976
	Proposed Estimator									
	(1)	0.063	0.258	0.266	0.087	0.300	0.312	0.071	0.404	0.410
	(2)	0.073	0.224	0.236	0.087	0.248	0.263	0.069	0.327	0.334
	(3)	0.064	0.225	0.234	0.093	0.250	0.267	0.071	0.296	0.304
	(4)	0.067	0.230	0.239	0.083	0.269	0.281	0.077	0.310	0.319
	(5)	0.061	0.253	0.260	0.078	0.295	0.305	0.057	0.348	0.352
	ATE for 2nd period									
	OLS	2.884	0.305	2.900	2.878	0.296	2.893	2.861	0.313	2.879
	FE	0.873	0.283	0.918	0.875	0.311	0.928	0.869	0.452	0.980
	Proposed Estimator									
	(1)	0.074	0.274	0.283	0.095	0.315	0.329	0.080	0.423	0.430
	(2)	0.083	0.243	0.256	0.090	0.274	0.288	0.078	0.345	0.354
	(3)	0.075	0.244	0.256	0.101	0.275	0.293	0.081	0.315	0.326
	(4)	0.074	0.248	0.259	0.091	0.290	0.304	0.087	0.336	0.347
	(5)	0.071	0.262	0.271	0.085	0.307	0.318	0.068	0.371	0.377
N=800	ATE for 1st period									
	OLS	2.874	0.228	2.883	2.876	0.242	2.886	2.870	0.234	2.880
	FE	0.870	0.208	0.894	0.874	0.253	0.910	0.886	0.362	0.957
	Proposed Estimator									
	(1)	0.066	0.214	0.224	0.067	0.257	0.266	0.074	0.325	0.333
	(2)	0.072	0.189	0.202	0.067	0.217	0.227	0.066	0.272	0.280
	(3)	0.070	0.184	0.196	0.079	0.204	0.219	0.069	0.242	0.252
	(4)	0.064	0.192	0.202	0.071	0.211	0.222	0.063	0.258	0.265
	(5)	0.067	0.218	0.228	0.063	0.246	0.253	0.056	0.302	0.307
	ATE for 2nd period									
	OLS	2.866	0.230	2.875	2.880	0.243	2.890	2.869	0.232	2.878
	FE	0.868	0.209	0.893	0.874	0.256	0.911	0.883	0.364	0.955
	Proposed Estimator									
	(1)	0.077	0.227	0.239	0.080	0.269	0.281	0.078	0.336	0.344
	(2)	0.081	0.205	0.220	0.072	0.241	0.251	0.070	0.287	0.296
	(3)	0.079	0.197	0.213	0.089	0.229	0.246	0.074	0.262	0.272
	(4)	0.072	0.207	0.219	0.077	0.236	0.248	0.066	0.277	0.285
	(5)	0.076	0.222	0.235	0.073	0.258	0.268	0.063	0.317	0.323

Note: Bias, SD, and RMSE are the mean bias, standard deviation, and the root mean squared error of the estimates across the simulations, respectively. For the Proposed Estimator, the rows (1), (2), (3), (4), and (5) report the results of the proposed estimator when Z_{it} is $\{W_{it}\}$, $\{X_{it,1}\}$, $\{X_{it,2}\}$, $\{X_{it,1}, X_{it,2}\}$, and $\{W_{it}, X_{it,1}, X_{it,2}\}$, respectively.

Table B.3: Monte Carlo Simulation Results: DGP3

		N=200			N=500			N=800		
	True Value	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
ATE for 1st period										
OLS	0.05	0.381	0.215	0.437	0.398	0.138	0.421	0.395	0.107	0.409
FE	0.05	-0.431	0.287	0.518	-0.418	0.180	0.455	-0.430	0.141	0.453
Proposed Estimator										
(1)	0.05	0.192	0.395	0.439	0.162	0.260	0.306	0.116	0.218	0.247
(2)	0.05	0.217	0.362	0.422	0.181	0.246	0.305	0.144	0.213	0.257
(3)	0.05	0.233	0.380	0.446	0.177	0.263	0.317	0.128	0.230	0.263
(4)	0.05	0.264	0.380	0.463	0.197	0.251	0.319	0.165	0.223	0.278
(5)	0.05	0.307	0.376	0.485	0.243	0.269	0.362	0.187	0.224	0.292
ATE for 2nd period										
OLS	1.05	0.393	0.220	0.450	0.400	0.136	0.422	0.387	0.107	0.401
FE	1.05	-0.423	0.291	0.513	-0.425	0.181	0.462	-0.431	0.139	0.453
Proposed Estimator										
(1)	1.05	0.041	0.433	0.434	0.059	0.276	0.282	0.040	0.231	0.234
(2)	1.05	0.101	0.347	0.362	0.110	0.234	0.258	0.081	0.202	0.218
(3)	1.05	0.153	0.430	0.456	0.144	0.286	0.320	0.105	0.254	0.274
(4)	1.05	0.153	0.373	0.403	0.142	0.244	0.282	0.122	0.221	0.252
(5)	1.05	0.153	0.363	0.394	0.161	0.251	0.298	0.124	0.217	0.250

Note: True Value is the true value of the ATE. Bias, SD, and RMSE are the mean bias, standard deviation, and the root mean squared error of the estimates across the simulations, respectively. For the Proposed Estimator, the rows (1), (2), (3), (4), and (5) report the results of the proposed estimator when Z_{it} is $\{W_{it}\}$, $\{X_{it,1}\}$, $\{X_{it,2}\}$, $\{X_{it,1}, X_{it,2}\}$, and $\{W_{it}, X_{it,1}, X_{it,2}\}$, respectively.

Table B.4: Monte Carlo Simulation Results: DGP4

		N=200			N=500			N=800		
	True Value	Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
ATE for 1st period										
OLS	0.01	0.338	0.215	0.401	0.355	0.138	0.381	0.352	0.108	0.368
FE	0.01	-0.455	0.290	0.539	-0.442	0.181	0.478	-0.454	0.142	0.476
Proposed Estimator										
(1)	0.01	0.191	0.664	0.690	0.130	0.450	0.468	0.066	0.369	0.374
(2)	0.01	0.186	0.658	0.684	0.128	0.430	0.449	0.080	0.344	0.353
(3)	0.01	0.173	0.653	0.675	0.120	0.462	0.477	0.084	0.370	0.379
(4)	0.01	0.217	0.642	0.677	0.124	0.453	0.469	0.077	0.359	0.367
(5)	0.01	0.251	0.631	0.678	0.158	0.449	0.476	0.093	0.368	0.379
ATE for 2nd period										
OLS	1.01	0.350	0.219	0.413	0.357	0.136	0.382	0.344	0.107	0.360
FE	1.01	-0.446	0.293	0.534	-0.449	0.183	0.485	-0.455	0.140	0.476
Proposed Estimator										
(1)	1.01	0.077	0.749	0.752	0.083	0.463	0.470	0.043	0.391	0.393
(2)	1.01	0.090	0.645	0.651	0.096	0.425	0.433	0.051	0.355	0.359
(3)	1.01	0.099	0.693	0.700	0.103	0.475	0.486	0.076	0.410	0.416
(4)	1.01	0.142	0.628	0.643	0.115	0.441	0.456	0.070	0.371	0.377
(5)	1.01	0.164	0.635	0.655	0.144	0.424	0.448	0.076	0.374	0.382

Note: True Value is the true value of the ATE. Bias, SD, and RMSE are the mean bias, standard deviation, and the root mean squared error of the estimates across the simulations, respectively. For the Proposed Estimator, the rows (1), (2), (3), (4), and (5) report the results of the proposed estimator when Z_{it} is $\{W_{it}\}$, $\{X_{it,1}\}$, $\{X_{it,2}\}$, $\{X_{it,1}, X_{it,2}\}$, and $\{W_{it}, X_{it,1}, X_{it,2}\}$, respectively.

Table B.5: Monte Carlo Simulation Results: Model 1

		T=10			T=20			T=30		
		Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
N=200	OLS	1.137	0.248	1.163	1.137	0.253	1.165	1.142	0.248	1.169
	FE	-0.005	0.211	0.211	-0.004	0.209	0.209	-0.003	0.208	0.208
	Proposed Estimator									
	(1)	0.011	0.117	0.122	-0.002	0.119	0.120	-0.006	0.105	0.106
	(2)	0.011	0.117	0.121	0.000	0.118	0.120	-0.005	0.104	0.105
	(3)	0.012	0.116	0.121	0.000	0.118	0.120	-0.005	0.105	0.106
	(4)	0.014	0.116	0.121	0.003	0.117	0.118	-0.004	0.104	0.105
	(5)	0.014	0.117	0.122	0.004	0.117	0.118	-0.003	0.104	0.105
N=500	OLS	1.138	0.159	1.149	1.140	0.159	1.151	1.142	0.157	1.153
	FE	-0.005	0.135	0.136	0.001	0.132	0.132	-0.001	0.131	0.131
	Proposed Estimator									
	(1)	-0.002	0.075	0.078	-0.004	0.072	0.074	-0.011	0.063	0.065
	(2)	-0.003	0.073	0.076	-0.004	0.072	0.074	-0.011	0.063	0.065
	(3)	-0.003	0.073	0.077	-0.004	0.072	0.074	-0.010	0.063	0.064
	(4)	-0.003	0.073	0.077	-0.004	0.071	0.073	-0.011	0.062	0.064
	(5)	-0.003	0.074	0.077	-0.003	0.071	0.073	-0.010	0.062	0.064
N=800	OLS	1.140	0.124	1.457	1.144	0.125	1.150	1.140	0.124	1.147
	FE	-0.001	0.105	0.105	-0.001	0.105	0.104	-0.001	0.104	0.104
	Proposed Estimator									
	(1)	-0.009	0.058	0.062	-0.010	0.057	0.059	-0.016	0.050	0.053
	(2)	-0.009	0.058	0.062	-0.010	0.057	0.059	-0.016	0.049	0.053
	(3)	-0.010	0.058	0.062	-0.009	0.057	0.059	-0.016	0.050	0.053
	(4)	-0.010	0.058	0.061	-0.009	0.056	0.058	-0.016	0.050	0.053
	(5)	-0.009	0.058	0.062	-0.010	0.056	0.058	-0.015	0.050	0.053

Note: Bias here is the time average of the means of the bias of the time-varying ATEs' estimates across the simulation. Similarly, SD and RMSE are the time averages of the standard deviations and the root mean squared error of the time-varying ATEs' estimates across the simulation. For the Proposed Estimator, the rows (1), (2), (3), (4), and (5) report the results of the proposed estimator when Z_{it} is $\{W_{it}\}$, $\{X_{it,1}\}$, $\{X_{it,2}\}$, $\{X_{it,1}, X_{it,2}\}$, and $\{W_{it}, X_{it,1}, X_{it,2}\}$, respectively.

Table B.6: Monte Carlo Simulation Results: Model 2

		T=10			T=20			T=30		
		Bias	SD	RMSE	Bias	SD	RMSE	Bias	SD	RMSE
N=200	OLS	2.803	0.435	2.836	2.790	0.432	2.824	2.797	0.438	2.831
	FE	0.936	0.290	0.980	0.939	0.287	0.981	0.942	0.284	0.984
	Proposed Estimator									
	(1)	-0.088	0.121	0.267	-0.052	0.106	0.175	-0.047	0.084	0.130
	(2)	-0.094	0.118	0.262	-0.050	0.107	0.175	-0.046	0.087	0.131
	(3)	-0.090	0.119	0.264	-0.050	0.107	0.176	-0.045	0.086	0.131
	(4)	-0.088	0.118	0.265	-0.049	0.105	0.174	-0.045	0.086	0.131
	(5)	-0.853	0.121	0.269	-0.049	0.104	0.174	-0.045	0.085	0.130
N=500	OLS	2.804	0.278	2.818	2.798	0.277	2.812	2.801	0.277	2.814
	FE	0.933	0.187	0.951	0.937	0.181	0.954	0.942	0.182	0.959
	Proposed Estimator									
	(1)	-0.106	0.076	0.220	-0.063	0.069	0.137	-0.055	0.057	0.103
	(2)	-0.107	0.073	0.217	-0.062	0.069	0.137	-0.054	0.058	0.104
	(3)	-0.106	0.074	0.218	-0.063	0.069	0.137	-0.054	0.058	0.104
	(4)	-0.107	0.074	0.218	-0.063	0.068	0.137	-0.054	0.058	0.104
	(5)	-0.105	0.076	0.220	-0.063	0.068	0.137	-0.054	0.058	0.104
N=800	OLS	2.803	0.221	2.812	2.797	0.215	2.805	2.801	0.218	2.809
	FE	0.929	0.147	0.940	0.935	0.143	0.946	0.941	0.143	0.951
	Proposed Estimator									
	(1)	-0.113	0.061	0.203	-0.069	0.057	0.125	-0.058	0.051	0.098
	(2)	-0.113	0.058	0.201	-0.068	0.057	0.125	-0.057	0.051	0.098
	(3)	-0.114	0.059	0.202	-0.068	0.057	0.125	-0.057	0.051	0.098
	(4)	-0.114	0.060	0.202	-0.069	0.056	0.125	-0.057	0.051	0.098
	(5)	-0.113	0.060	0.203	-0.069	0.056	0.125	-0.057	0.050	0.098

Note: Bias here is the time average of the means of the bias of the time-varying ATEs' estimates across the simulation. Similarly, SD and RMSE are the time averages of the standard deviations and the root mean squared error of the time-varying ATEs' estimates across the simulation. For the Proposed Estimator, rows (1), (2), (3), (4), and (5) report the results of the proposed estimator when Z_{it} is $\{W_{it}\}$, $\{X_{it,1}\}$, $\{X_{it,2}\}$, $\{X_{it,1}, X_{it,2}\}$, and $\{W_{it}, X_{it,1}, X_{it,2}\}$, respectively.