Estimation of time-varying average treatment effects using panel data when unobserved fixed effects affect potential outcomes differently

Shosei Sakaguchi* †

July 10, 2016

Abstract

This paper proposes a new approach to estimate the time-varying average treatment effect using panel data to control for unobserved fixed effects. The approach allows identifying the average treatment effect on the entire population, even if the fixed effects affect potential outcomes under treatment and no treatment differently, which can cause heterogeneity in treatment effects among unobserved characteristics. Note that a popularly used standard difference-in-differences approach can only identify the average treatment effect on the treated. Moreover, the proposed approach allows time-varying treatment effects. The approach exploits panel data with a specific structure in which the treatment exposure expands to the entire population over time. I apply the proposed approach to estimate the effect of the introduction of electronic voting technology for the reduction of residual votes in Brazilian elections.

Keywords: Program evaluation, Panel data, Fixed effects, Difference-in-Differences, Average treatment effects
JEL classification: C21, C23

*Graduate School of Economics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan. Tel: +81 75 753 7102; fax: +81 75 753 7193. Email: sakaguchi.shosei.75x@st.kyoto-u.ac.jp.

†Research Fellow, Japan Society for the Promotion of Science.

1 Introduction

This paper proposes a new approach to estimate the time-varying average treatment effect (ATE) using panel data to control for unobserved time-invariant confounders (i.e., fixed effects). It allows identification of the average treatment effect on the entire population, even if the fixed effects have different influences on potential outcomes under treatment and no treatment, which can cause heterogeneity in treatment effects among unobserved characteristics. Moreover, it allows for time-varying treatment effects.

A fixed-effects (FE) estimation and a difference-in-differences (DID) estimation are popularly used to control unobserved fixed effects; however, they cannot consistently estimate the time-varying ATE for the entire population when unobserved fixed effects affect the two potential outcomes differently (Wooldridge, 2010, Ch. 21; Lechner, 2011). The FE estimation requires the same influence from unobserved fixed effects, a very restrictive assumption because there are many cases where unobserved fixed effects have varying influences on the two potential outcomes. For example, in job training program evaluations, an individual's unobserved potential ability should have a different influence on wages whether or not the one participates in the training. On the other hand, the standard DID estimation consistently estimates the ATE on the treated but not on the entire population, regardless of whether unobserved fixed effects have different influences on the two potential outcomes (Lechner, 2011). The proposed approach can estimate the time-varying ATE on the entire population in this situation.

To identify the ATE, the proposed approach exploits uniquely structured panel data wherein the treatment exposure expands from no units to all units across time. Many policy interventions examined in empirical studies of program evaluation have this property. Such policies, for instance, include reforms to the no-fault unilateral divorce laws in the U.S. (Wolfers, 2006), reforms to the federally mandated sick leave insurance scheme in Germany (Puhani and Sonderhof, 2010), and the introduction of electronic voting (EV) technology in Brazil (Fujiwara, 2015). By exploiting this panel data structure, this paper proposes an approach to consistently estimate the time-varying ATE, even when unobserved fixed effects affect the two potential outcomes differently.

The outline of the paper is as follows. Section 2 describes the setting and the model specification. Section 3 proposes an estimation approach and derives its asymptotic properties. Section 4 describes its empirical application. Section 5 concludes.

2 Setting and Model Specification

In this paper, we suppose that a sample $\{Y_{it}, D_{it}, X_{it}\}$ is observed for N units $i = (1, 2, \dots, N)$ across T time periods $t = (1, 2, \dots, T)$, where $T \geq 3$. We assume N is large while T is small. Y_{it} denotes an observed outcome. X_{it} denotes a $K \times 1$ vector of observed covariates that may include a time dummy, observed time-varying confounders, and interactions of the time dummy and observed time-invariant confounders. Each unit is grouped by $D_{it} \in \{0, 1\}$ such that $D_{it} = 1$ indicates treatment in period t.

Our interest is in the ATE on the entire population for each intermediate period $t = 2, \dots, T-1$:

$$\tau_t^{ate} = E[Y_{it}(1) - Y_{it}(0)],$$

where $Y_{it}(1)$ and $Y_{it}(0)$ denotes two potential outcomes under treatment and no treatment for unit i in period t, respectively. The observed outcome Y_{it} is expressed as $Y_{it} = D_{it}Y_{it}(1) + (1 - D_{it})Y_{it}(0)$.

Given X_{it} and unobserved fixed effects C_i, we assume the potential outcomes are:

$$Y_{it}(0) = X_{it}'\beta_t^0 + g^0(C_i) + u_{it}^0, \tag{1}$$

$$Y_{it}(1) = X_{it}'\beta_t^1 + g^1(C_i) + u_{it}^1, \tag{2}$$

where u_{it}^j is a mean-zero error term defined as $u_{it}^j = Y_{it}(j) - E[Y_{it}(j) \mid X_{it}, C_i]$ for $j = 0, 1$. The coefficients of observed covariates, β_t^0 and β_t^1, may be time varying. $g^0(C_i)$ and $g^1(C_i)$ are possibly different functions of C_i, which implies that unobserved fixed effects may have different influences on the two potential outcomes. We can also suppose a nonlinear form of the observed covariates. Note that the assumed potential outcome models do not allow for interactions of $g^0(C_i)$ and $g^1(C_i)$ with X_{it}, and rule out influences of past treatments on future outcomes unlike the dynamic potential outcome framework suggested by, for example, Lechner and Miquel (2009).

In the potential outcome models (1) and (2), τ_t^{ate} is expressed as

$$\tau_t^{ate} = E[X_{it}]'(\beta_t^1 - \beta_t^0) + E[g^1(C_i) - g^0(C_i)]. \tag{3}$$

As seen in Equation (3), the difference between $g^0(C_i)$ and $g^1(C_i)$ is a cause of heterogeneous treatment effects among unobserved fixed effects. The time variation of observed covariates and their coefficients cause the ATE time variation.

Although it is natural to consider that unobserved fixed effects have different influences on the potential outcomes in many cases, the popular FE estimation does not provide consistent estimations of τ_t^{ate} when $g^0(C_i)$ and $g^1(C_i)$ are different, but rather require that they are the same. The standard DID estimation provides consistent estimation of the ATE on the treated rather than τ_t^{ate}. The next section proposes a new approach to estimate τ_t^{ate} in the supposed potential outcome models.

For the defined variables, we impose the following assumptions.

Assumption 1. $\{\{Y_{it}, D_{it}, X_{it}\}_{t=1}^T, C_i\}$ are i.i.d. across i.

Assumption 2. $E[Y_{it}(j) \mid D_{it}, X_{i1}, \ldots, X_{iT}, C_i] = E[Y_{it}(j) \mid X_{it}, C_i]$ for all $j = 0, 1$ and t.

Assumption 3. $D_{i1} = 0$ and $D_{iT} = 1$ for all i.

Assumption 4. $E[(1-D_{it})\cdot(X_{it}', -X_{i1}')'(X_{it}', -X_{i1}')]$ and $E[D_{it}\cdot(X_{it}', -X_{iT}')'(X_{it}', -X_{iT}')]$ are nonsingular.

Assumption 2 corresponds to the required assumption for the FE estimation and has two meanings. The one is that the treatment assignment and the potential outcomes are mean independent conditioning on the observed covariates and unobserved fixed effects. This assumption is weaker than the normal conditional mean independence assumption that requires only observed covariates as sufficient for independence. The other is that the potential outcomes in period t do not depend on past and future observed covariates, which means strict exogeneity known in the panel data literature. Assumption 3 is crucial for the proposed approach, and means that no units are treated in the initial period 1, while all units are treated in the final period T. Under Assumption 3, the panel data is structured such that treatment exposure ranges from no units to all units across time. Assumption 4 rules out time-invariant covariates in X_{it}

Note that, under the potential outcome models (1) and (2), Assumption 2 implies common trend assumptions for the mean potential outcome under no treatment and, moreover, for the mean potential outcome under treatment. The proposed method is based on both of these common trend assumptions. Considering this point, Assumption 2 is stronger than the conventional common trend assumption for the standard DID estimation since it imposes a common trend assumption only for the mean potential outcome under no treatment. I discuss these points in more detail in the supplementary appendix.

3 Estimation

Given the above setting and model specification, this section proposes a new approach to estimate τ_t^{ate} consistently for each period $t = 2, \ldots, T-1$. Recall that the panel data we consider have a structure in which the treatment exposure expands from no units to all units across time. By exploiting this structure, the proposed approach allows for consistent estimates of $E[Y_{it}(0)]$ using data for periods t and 1, and $E[Y_{it}(1)]$ using data for periods t and T, controlling unobserved fixed effects. Then, τ_t^{ate} can be consistently estimated. By using data for different periods to estimate $E[Y_{it}(0)]$ and $E[Y_{it}(1)]$, we can control unobserved fixed effects, even when they have different influences on the two potential outcomes.

The approach consists of the following three steps.

First step:

For the potential outcome under no treatment, we consider the difference model for periods t and 1;

$$Y_{it}(0) - Y_{i1}(0) = X'_{it}\beta_t^0 - X'_{i1}\beta_1^0 + u_{it}^0 - u_{i1}^0, \tag{4}$$

where the component of unobserved fixed effects $g^0(C_i)$ is differenced out. Using a subsample of units with $D_{i1} = D_{it} = 0$, we obtain the OLS estimators of β_t^0 and β_1^0 denoted by $\hat{\beta}_t^0$ and $\hat{\beta}_1^0$, respectively.

Then, using the full sample, we estimate $E[Y_{it}(0)]$ as follows:

$$\widehat{E[Y_{it}(0)]} = \frac{1}{N}\sum_{i=1}^{N}[X'_{it}\hat{\beta}_t^0 - X'_{i1}\hat{\beta}_1^0 + Y_{i1}]. \tag{5}$$

Second step:

The second step is a symmetry of the first step. For the potential outcome under treatment, we consider the following difference model for periods t and T:

$$Y_{it}(1) - Y_{iT}(1) = X'_{it}\beta_t^1 - X'_{iT}\beta_T^1 + u_{it}^1 - u_{iT}^1. \tag{6}$$

Using the subsample of units with $D_{it} = D_{iT} = 1$, we obtain the OLS estimator of β^1 denoted by $\hat{\beta}^1$.

Then, using the full sample, we estimate $E[Y_{it}(1)]$ as follows:

$$\widehat{E[Y_{it}(1)]} = \frac{1}{N}\sum_{i=1}^{N}[X_{it}\hat{\beta}_t^1 - X'_{iT}\hat{\beta}_T^1 + Y_{iT}]. \tag{7}$$

Third step:

Finally, τ_t^{ate} is estimated as follows:

$$\hat{\tau}_t^{ate} = \widehat{E[Y_{it}(1)]} - \widehat{E[Y_{it}(0)]}. \tag{8}$$

The estimator from the above sequential steps is included in the class of GMM estimators, and its asymptotic property can be derived from GMM theory (Newey, 1984). The estimator $\hat{\tau}_t^{ate}$ is consistent and asymptotically normal under Assumptions 1-4 and some regularity conditions.

Proposition 1. Under Assumptions 1-4 and Assumption A.1 in the appendix, $\hat{\tau}_t^{ate} \xrightarrow{p} \tau_t^{ate}$ and $\sqrt{N}(\hat{\tau}_t^{ate} - \tau_t^{ate}) \xrightarrow{d} N(0, V_t)$ where:

$$\begin{aligned} V_t =& E[\{(X_{it}'\beta_t^1 - X_{iT}'\beta_T^1 + Y_{iT}) - (X_{it}'\beta_t^0 - X_{i1}'\beta_1^0 + Y_{i1}) - \tau_t^{ate} \\ &- E[X_{it}', -X_{i1}']'E\left[(1-D_{it}) \cdot (X_{it}', -X_{i1}')'(X_{it}', -X_{i1}')\right]^{-1}\left[(1-D_{it}) \cdot (X_{it}', -X_{i1}')'(u_{it}^0 - u_{i1}^0)\right] \\ &+ E[X_{it}', -X_{iT}']'E\left[D_{it} \cdot (X_{it}', -X_{iT}')'(X_{it}', -X_{iT}')\right]^{-1}\left[D_{it} \cdot (X_{it}', -X_{iT}')'(u_{it}^1 - u_{iT}^1)\right]\}^2]. \end{aligned} \quad (9)$$

The appendix provides the proof and the regularity conditions. Given $\hat{\beta}_1^0$, $\hat{\beta}_t^0$, $\hat{\beta}_t^1$, $\hat{\beta}_T^1$, and $\hat{\tau}_t^{ate}$, the asymptotic variance V_t is consistently estimated as a sample analogue of (9).

4 Empirical Application

I apply the proposed approach to estimate the introduction of EV technology for the reduction of residual votes in Brazilian elections. Brazil holds elections for state officials every four years, and all states have the same election date. The 1994 elections used only paper ballots in all municipalities; in the 1998 elections, municipalities with more than 40,500 registered voters used EV technology, while municipalities below this threshold used paper ballots. In 2002, all municipalities used EV.

I analyze the data used in Fujiwara (2015). He estimates the effect of EV technology by exploiting a regression discontinuity design (RDD) embedded in the introduction of EV. While he estimates the ATE for a municipality with 40,500 voters, I estimate the ATE for the entire population by applying the proposed approach.

Table 1 presents the estimates using the proposed approach, DID, and Fujiwara's (2015) RDD. The parameter of interest is the ATE on valid votes as a share of turnout in the 1998 elections. In applying the proposed approach, models (1) and (2) include year dummies, interaction terms of year dummies with the state's fixed effects, and the municipality's fixed effects. In the DID estimation, the restriction

$g^1(C_i) = g^0(C_i)$ is imposed for the municipality's fixed effects and the ATE on the treated is estimated. Fujiwara's (2015) RDD estimation controls the state's fixed effects. The result shows that the proposed approach yields a slightly higher estimate than the DID and RDD estimates but it is not so quantitatively different from the others. In this sense, the result indicates that Fujiwara's (2015) RDD result is robust for the entire population.

Table 1: Estimates of ATE for the introduction of EV technology in Brazil[a]

	Proposed approach	DID	RDD
Valid Votes/Turnout (1998 Election)	0.152	0.146	0.139
	(0.001)	(0.003)	(0.013)
$[N1, N0]$	$[307, 4502]$	$[307, 4502]$	$[91, 174]$

[a] Robust standard errors are in parentheses. $N1$ and $N0$ indicate the number of municipalities in the treated and untreated groups, respectively. For RDD estimation details, see Fujiwara (2015).

5 Conclusion

This paper proposes a new approach to estimate the time-varying ATE on the entire population using panel data with a specific structure in which the treatment exposure expands to the entire population over time. This data structure enables consistent estimations using the proposed approach, even when unobserved fixed effects affect the two potential outcomes differently. While the main text of the paper focuses only on the panel data setting, the proposed method can be similarly constructed for repeated cross-section data, which is discussed thoroughly in the supplementary appendix.

Acknowledgments

I am grateful to my supervisors Yoshihiko Nishiyama and Ryo Okui, Ken Yamada, and an anonymous referee for their helpful comments and suggestions. This work was partially supported by JSPS KAKENHI Grant Number JP16J01170.

References

Fujiwara, T., 2015. Voting technology, political responsiveness, and infant health: Evidence from Brazil. Econometrica 83, 423-464.

Lechner, M., 2011. The estimation of causal effects by difference-in-difference methods. Foundations and Trends in Econometrics 4, 165-224.

Lechner, M., Miquel, R., 2010. Identification of the effects of dynamic treatments by sequential conditional independence assumptions. Empirical Econ. 39, 111-137.

Newey, W.K., 1984. A method of moments interpretation of sequential estimators. Economics Letters 14, 201-206.

Puhani, P.A., Sonderhof, K., 2010. The effects of a sick pay reform on absence and on health-related outcomes. J. Health Economics 29, 285-302.

Wolfers, J., 2006. Did unilateral divorce laws raise divorce rates? A reconciliation and new results. Amer. Econ. Rev. 96, 1802-1820.

Wooldridge, J.M., 2010. Econometric analysis of cross section and panel data, 2nd edition. The MIT Press.