Economics Letters 220 (2022) 110874

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Synthetic difference-in-differences estimation with staggered treatment timing

Zachary Porreca

Department of Economics, West Virginia University, Morgantown, WV, United States of America

ARTICLE INFO

Article history:
Received 28 March 2022
Received in revised form 23 August 2022
Accepted 21 September 2022
Available online 27 September 2022

Keywords:
Synthetic difference-in-differences
Staggered treatment
Difference-in-differences
Causal identification

ABSTRACT

This note formalizes the synthetic difference-in-differences estimator for staggered treatment adoption settings, as briefly described in Arkhangelsky et al. (2021). To illustrate the importance of this estimator, I use replication data from Abrams (2012). I compare the estimators obtained using SynthDiD, TWFE, the group time average treatment effect estimator of Callaway and Sant'Anna (2021), and the partially pooled synthetic control method estimator of Ben-Michael et al. (2021) in a staggered treatment adoption setting. I find that in this staggered treatment setting, SynthDiD provides a numerically different estimate of the average treatment effect. Simulation results show that these differences may be attributable to the underlying data generating process more closely mirroring that of the latent factor model assumed for SynthDiD than that of additive fixed effects assumed under traditional difference-in-differences frameworks.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Causal inference has become the dominant aim of empirical micro-economic research, making unbiased causal estimation an issue at the forefront of the econometric literature. Difference-in-differences (DiD) estimation is now applied in settings with staggered treatment adoption, rather than solely in the more restrictive simultaneous treatment adoption framework. Recent research has focused on best practices in these settings (De Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2021; Roth, 2022; Goodman-Bacon, 2021; Baker et al., 2021; Callaway and Sant'Anna, 2021). Simultaneously, an alternative path to estimating causal inference, the Synthetic Control Method (SCM) of Abadie and Gardeazabal (2003) and Abadie et al. (2010), has been developed. This method synthesizes an optimal control group for settings with only one treated unit, through the application of weights to pre-treatment period control units. SCM allows researchers to make use of empirical causal inference in case study settings. This method has been extended to settings with multiple treated units. Ben-Michael et al. (2021) explored these multi-unit synthetic control estimators and began the formalization of pooled synthetic control estimators.

Arkhangelsky et al. (2021) introduced a new causal inference estimator that can be employed in both traditional DiD settings and in those settings typically in the domain of SCM, such as those with a single treated unit and no clear control group. Their method, synthetic difference-in-differences estimation (SynthDiD), does not rely on parallel trends assumptions or assumptions of exogeneity of the treatment as in DiD. SynthDiD does this by introducing the weighting of both pre-treatment time periods and cross-sectional units in the construction of a synthetic counterfactual for causal estimation. This method allows more heterogeneity of outcomes and has been suggested to improve the precision of the estimator (Arkhangelsky et al., 2021). Details of this method are discussed in the next section. The application of this estimator to staggered treatment timing settings has yet to be thoroughly explored.

The process by which a SynthDiD estimator is obtained by weighting of individual cohort estimators is briefly described in an appendix to Arkhangelsky et al. (2021). However, the actual functional form of the estimator has not been formalized until the present note. Here, I offer a formalization of this estimator, present an estimator of its variance, and offer a practical example demonstrating its implementation. To illustrate this estimator's importance and distinction, I employ data from Abrams (2012a,b) to demonstrate the differences in estimators obtained through SynthDiD, a standard two-way fixed effects DiD estimator, the group-time average treatment effect estimator of Callaway and Sant'Anna (2021), and the partially pooled synthetic control method estimator of Ben-Michael et al. (2021). This paper is meant to serve as an introduction to the use of SynthDiD estimation in the staggered treatment timing settings often encountered in empirical research.

E-mail address: zjp00003@mix.wvu.edu.

https://doi.org/10.1016/j.econlet.2022.110874
0165-1765/© 2022 Elsevier B.V. All rights reserved.

First, I will present a brief overview of the single treatment period SynthDiD estimator as described by Arkhangelsky et al. (2021). I then proceed to formalize the staggered treatment timing SynthDiD estimator and describe the practicalities of its estimation. Last, I will demonstrate the outcomes obtained by implementing this estimator using open source software.

2. Synthetic DiD

In depth detail regarding the performance, precision, and theoretical basis of the synthetic difference-in-differences estimator can be found in Arkhangelsky et al. (2021). In short, this estimator is novel in its inclusion of both time and unit weights to create a reliable counterfactual, its inclusion of a unit fixed effect (which is missing in SCM), and its reliance on weighting control observations to achieve parallel trends with treated observations rather than weighting control observations to match with treated observations directly.

The synthetic difference-in-differences estimator assumes that the data generating process follows a latent factor model:

$$\boldsymbol{Y} = \boldsymbol{L} + (\boldsymbol{W\tau})_{i,t} + \boldsymbol{E} \quad where \quad \boldsymbol{L} = \boldsymbol{\Gamma\Upsilon^T} \quad and \quad (\boldsymbol{W\tau})_{i,t} = \boldsymbol{W_{i,t}\tau_{i,t}} \tag{1}$$

In the above data generating process, $W_{i,t}$ represents a binary treatment indicator that is equal to 1 if unit i is treated in period t and 0 otherwise. Γ is a vector of latent time factors and Υ is a vector of latent unit factors. τ is the average treatment effect and E is an error matrix.

The average treatment effect in this data generating process is defined as:

$$\tau = \frac{1}{N_{tr}T_{post}} \sum_{i=N_{co}+1}^{N} \sum_{t=T_{pre}+1}^{T} \tau_{it} \tag{2}$$

N_{tr} and N_{co} represent the number of treatment and control group observations respectively. T_{pre} and T_{post} represent the number of time periods before and after treatment.

The estimator of the average treatment effect, $\hat{\tau}$, is formally defined as follows:

$$\hat{\tau} = \left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \hat{\delta}_i\right) - \left(\sum_{i=1}^{N_{co}} \hat{\omega}_i \hat{\delta}_i\right) \tag{3}$$

$$\hat{\delta}_i = \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it}\right) - \left(\sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it}\right) \tag{4}$$

Y_{it} is the observed outcome variable of interest. $\hat{\lambda}_t$ and $\hat{\omega}_i$, the time period and unit weights, are described below. Together, the overall estimator can be expressed as follows:

$$\hat{\tau} = \left[\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it}\right)\right] - \left[\sum_{i=1}^{N_{co}} \hat{\omega}_i \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it}\right)\right] \tag{5}$$

The procedure by which the weights $\hat{\lambda}_t$ and $\hat{\omega}_i$ are chosen is described in Arkhangelsky et al. (2021). The time period weights are defined as follows:

$$\begin{aligned}
&(\hat{\lambda}_0, \hat{\lambda}^{sdid}) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\operatorname{argmin}} \ \ell_{time}(\lambda_0, \lambda) \ \text{ where} \\
&\ell_{time}(\lambda_0, \lambda) = \sum_{i=1}^{N_{co}} (\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it})^2, \\
&\Lambda = \{\lambda \in \mathbb{R}_+^T : \sum_{t=1}^{T_{pre}} \lambda_t = 1, \lambda_t \\
&\quad = T_{post}^{-1} \ \text{ for all } \ t = T_{pre} + 1, \ldots, T\}
\end{aligned} \tag{6}$$

Similarly, the unit weights are defined as follows:

$$\begin{aligned}
&(\hat{\omega}_0, \hat{\omega}^{sdid}) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\operatorname{argmin}} \ \ell_{unit}(\omega_0, \omega) \ \text{ where} \\
&\ell_{unit}(\omega_0, \omega) = \sum_{t=1}^{T_{pre}} (\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it})^2 + \zeta T_{pre} ||\omega||_2^2, \\
&\Omega = \{\omega \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} \omega_i \\
&\quad = 1, \omega_i = N_{tr}^{-1} \ \text{ for all } \ i = Nco + 1, \ldots, N\},
\end{aligned} \tag{7}$$

The regularization parameter, ζ, is defined as:

$$\begin{aligned}
&\zeta = (N_{tr}T_{post})^{1/4}\hat{\sigma} \ \text{ with } \ \hat{\sigma}^2 = \frac{1}{N_{co}(T_{pre}-1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} (\Delta_{it} - \bar{\Delta})^2, \\
&\text{where } \ \Delta_{it} = Y_{i(t+1)} - Y_{it}, \ \text{ and } \ \bar{\Delta} = \frac{1}{N_{co}(T_{pre}-1)} \sum_{i=1}^{N_{co}} \sum_{t=1}^{T_{pre}-1} \Delta_{it}
\end{aligned} \tag{8}$$

3. Staggered treatment timing

Arkhangelsky et al. (2021) describe in their appendix a procedure for obtaining a weighted average of each treatment cohort's individual $\hat{\tau}$ to estimate the aggregate group average treatment effect. This proposed estimator is obtained by iterating through the SynthDiD algorithm on subsets of the data limited to the entire never-treated subset of units and each treatment cohort individually. This method is conceptually similar to that of Ben-Michael et al. (2021) and Callaway and Sant'Anna (2021).

Assuming the same data generating process described in Arkhangelsky et al. (2021) (and shown in Eq. (1)), the average treatment effect in the staggered treatment setting is defined as follows:

$$\tau = \frac{1}{N_{tr}T_{post}L} \sum_{\ell \in L | \ell > 0}^{L} \sum_{i=N_{co}+1}^{N} \sum_{t=T_{pre}+1}^{T} \tau_{it\ell} \tag{9}$$

The subscript $\ell \in L : \{0, 1, \ldots . \ell\}$ indexes a cohort membership. $\ell = 0$ is the cohort of never-treated control group observations.

This parameter can be estimated by allowing for a distinct $\hat{\tau}_\ell$ for each of L treatment cohorts.

$$\hat{\tau}_\ell = \left[\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it\ell} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it\ell}\right)\right] - \left[\sum_{i=1}^{N_{co}} \hat{\omega}_i \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it}\right)\right] \tag{10}$$

$\forall \ell \in L$. These individual $\hat{\tau}_\ell$ estimators are combined together by weighted averaging. The weights, μ_ℓ, are equal to the proportion of treated units that belong to each cohort, ℓ (Arkhangelsky et al., 2021). As such, the staggered treatment timing SynthDiD estimator can be formally expressed as:

$$\hat{\tau} = \sum_{\ell \in L | \ell > 0}^{L} \left(\mu_\ell \cdot \hat{\tau}_\ell \right) \tag{11}$$

The weight applied to the ℓth $\hat{\tau}$ cohort estimator is μ_ℓ. This weight is defined as:

$$\mu_\ell = \frac{N_\ell}{\sum_{\ell \in L | \ell > 0} N_\ell} \tag{12}$$

This weight is equivalent to the proportion of the of non-zero row-sum rows in the treatment assignment matrix **W** for which the first non-zero column of that row is column ℓ.[1]

Applying these weights, the overall $\hat{\tau}$ estimator can be expressed as:

$$\begin{aligned} \hat{\tau} = \sum_{\ell \in L | \ell > 0}^{L} \Bigg(& \left[\frac{N_\ell}{\sum_{\ell \in L | \ell > 0} N_\ell} \right] \\ & \times \left[\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it\ell} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it\ell} \right) \right] \\ & - \left[\sum_{i=1}^{N_{co}} \hat{\omega}_i \left(\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \right) \right] \Bigg) \end{aligned} \tag{13}$$

Practically, this estimator is simply a weighted average of cohort-specific estimated average treatment effects, where the weight applied to any individual cohort's specific estimator is equal to the proportion of treatment group observations that originate in a specific cohort Arkhangelsky et al. (2021).

Following Kahn (2015), it is a property of the influence functions of such estimators of summary parameters that the following holds (with $N = \sum_{\ell \in L | \ell > 0}^{L} N_\ell$):

$$\frac{1}{\sqrt{N}} (\hat{\tau} - \tau) = \sum_{i=1}^{N} \psi_\tau(x_i) + o_p(1) \tag{14}$$

Where $\psi_\tau(x_i)$ is the influence function for the nth observation and the summary parameter, τ. As such, the variance of the summary parameter $\hat{\tau}$ can be computed through the procedure described in Erickson and Whited (2002) and Kahn (2015). Paraphrasing Kahn (2015), this procedure entails calculating the empirical equivalents of the influence functions for each estimator and stacking them into a single matrix, Ψ, in which the rows correspond to each estimator and the columns to each observation. This procedure is similar to that used to calculate the variance of the summary parameter in Callaway and Sant'Anna (2021). The variance–covariance matrix for the individual cohort estimators can be calculated as follows:

$$\hat{V} = \frac{1}{N^2} (\Psi^T \Psi) \tag{15}$$

Thus, using the $\ell \times 1$ vector of weights, μ, described in Eq. (12), an estimator for the variance of the aggregated summary parameter can be computed as follows:

$$\hat{V}_{\hat{\tau}} = \mu^T \hat{V} \mu \tag{16}$$

[1] These weights correspond to those described in the appendix to Arkhangelsky et al. (2021).

4. Application

To demonstrate the use of the staggered-treatment timing synthetic difference-in-differences estimator, replication data from Abrams (2012a,b) are used to estimate the causal impact of a law adopted by different states in different years. This estimation is done with a typical two-way fixed effects OLS estimator (TWFE), the synthetic difference-in-differences estimator presented here, the partially-pooled synthetic control estimator of Ben-Michael et al. (2021), and an aggregated group-time average treatment effect estimator (CS) obtained through the estimation procedure detailed in Callaway and Sant'Anna (2021).[2]

Abrams (2012a,b) estimates the deterrent effect of firearm sentencing enhancements. Rather than systematically replicating the entirety of Abrams (2012a,b), replication data aggregated at the state level (used in an appendix to that paper) is used here for demonstrative purposes. Further, for simplicity of computation and demonstration, data is limited to a balanced panel of 45 states measured in each of 38 years.

$$y_{it} = \tau \cdot W_{it} + \gamma_i \upsilon_t^T + \epsilon_{it} \tag{17}$$

This equation is analogous to the latent factor model described as the data generating process in Eq. (1).[3] For the analysis herein, $\hat{\tau}$ is the parameter of interest: the average treatment effect. W_{it} is an indicator equal to one if a state i is treated in year t. Once a unit is treated, it remains treated for all subsequent time periods. The outcome variable of interest is the per-capita armed robbery rate in state i in year t. The treatment is the enactment of a state-wide sentencing enhancement for crimes that involve firearms.

First, I evaluate whether there is a violation of the parallel trends assumption that the traditional difference-in-differences estimation relies on for identification. An event study is estimated following the example of He and Wang (2017). Periods greater than 5 periods from treatment are aggregated (to ± 6 on either tail). Further, $t = -1$ is omitted to serve as the point of reference.

$$y_{it} = \sum_{k > -5, k \neq -1}^{k > 6} W_{it}^k \cdot \delta_k + \beta_t + \alpha_i + \epsilon_{it} \tag{18}$$

W_{it}^k is an indicator equal to one if state i in year t is k years away from initially enacting a sentencing enhancement. δ_k is the k period specific event-study estimator. The other variables and parameters remain unchanged. Eq. (18) assumes the additive form of the latent factor model in the data generating process discussed in Footnote 4. By making the assumption that $\gamma_i \upsilon_t^T = \lambda_i + \omega_t$, this event study equation is reconciled with the assumed data generating process.

The results of this estimation are displayed below. Given that some non-zero trend is evident in the pre-treatment periods, inference relying on the parallel trends assumption of difference-in-differences may be unreliable. A Wald χ^2 test of the joint significance of the pre-treatment estimators can serve as a parametric test of that parallel trends assumption (Roth, 2022). This test yields a χ^2 value of 2.01 with p-value of 0.075. As such, the null-hypothesis of joint significance fails to be rejected at the 5% level. Further analysis of this assumption is beyond the scope of this paper. However, there is some evidence that this assumption may be violated here, introducing unwarranted bias into the fixed-effects estimator (see Fig. 1).

[2] The individual cohort estimators for the group-time average treatment effect estimator (CS) are estimated using the doubly robust estimation procedure as detailed in Callaway and Sant'Anna (2021).

[3] As mentioned in Arkhangelsky et al. (2021), if the latent factor component of the data generating process takes an additive form ($L_{it} = \alpha_i + \beta_t$) this is analogous to the standard two-way fixed effects estimator utilized in most difference-in-differences frameworks.

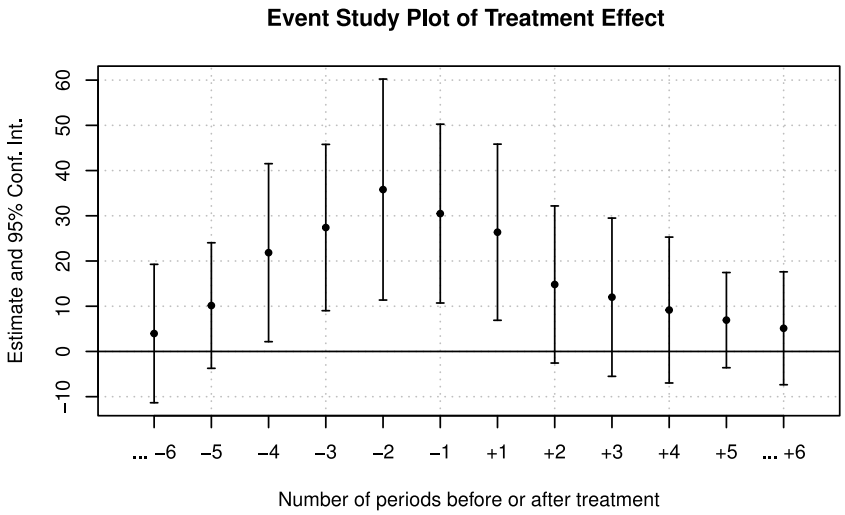

Fig. 1. Observed treatment effect by number of years pre or post treatment.

Table 1
Estimation of $\hat{\tau}$ without covariates.

	Dependent variable:			
	Per-capita armed robbery rate			
	SynthDiD	TWFE	CS	Partially-pooled SCM
$\hat{\tau}$	−16.697***	−20.182***	−23.724**	−14.095
	(0.364)	(7.430)	(5.243)	(10.177)
Observations		1710		1520

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are in parentheses. Standard errors for the TWFE estimator are clustered at the state level. Standard errors for the SynthDiD estimator are calculated as described in Eqs. (14)–(16). Standard errors for the Callaway and Sant'Anna estimator are bootstrapped. Standard errors for the partially-pooled SCM estimator were constructed using a jackknife procedure. For the Partially-pooled SCM estimator, observations for units that were already treated in the initial period are dropped, as the estimator's usage requires.

Estimating $\hat{\tau}$ under the aforementioned specifications without covariates yields the following Table 1.

SynthDiD produces an estimate of a smaller magnitude than both TWFE and CS. The variance of this estimator is smaller than all alternatives. This evidences the precision and reliability of the SynthDiD estimator discussed in Arkhangelsky et al. (2021).

In this applied context: while the impact of the sentencing enhancement is consistently estimated here to reduce armed robbery rates, SynthDiD estimates the effect (in the absence of covariates[4]) to be markedly more conservative than the alternative approaches. A brief discussion of what may be causing the discrepancy between these point estimates is included with the simulation results.

5. Simulation results

A simple simulation is useful in evaluating the performance of this SynthDiD estimator and provides some insight into the differences in point estimates from the empirical results. 1000 simulated datasets were generated with their parameters calibrated to those of the Abrams (2012a,b) replication data. Treatment is randomly assigned, with the true value of τ for each treated unit being normally distributed around −20 with a standard deviation of 2. The performance of the four estimators considered previously are compared in Table 2. The first row (with $\hat{\tau}_{additiveFE}$) estimates the $\hat{\tau}$ average treatment effect parameter on data simulated to follow an additive fixed effects data generating process, while the second row (with $\hat{\tau}_{Factor}$) estimates the same parameter on data simulated to follow a latent factor model data generating process.

This simple simulation exercise yields similar point estimates and standard errors for both SynthDiD and TWFE.[5] This is unsurprising, as the data generating process of the first series of simulations relies on the assumption of an additive form for its latent factor component (see Footnote 4). When this is the case, typical two-way fixed effects DiD will consistently recover the parameter τ (Arkhangelsky et al., 2021). SynthDiD has performed almost identically to TWFE in this context. This provides some explanation for the difference in point estimates seen in the above empirical section. It implies that the data generating process underlying the replication data utilized follows an interactive fixed effects model rather than the additive form TWFE relies on.[6] The second specification, in which the data generating process follows

[4] SynthDiD estimation (including in the staggered treatment timing setting) allows for the potential inclusion of covariates. As detailed in a footnote to Arkhangelsky et al. (2021), an adjusted y_{it}^{R} outcome variable is the residual from regressing y_{it} on the matrix of covariates:

$$y_{it}^{R} = y_{it} - X_{it}\hat{\beta} \tag{19}$$

The systematic effect of unit-specific time-varying X_{it} covariates on y are partialled out as in OLS, providing an adjusted outcome variable: y_{it}^{R}. This is the component of y that remains unexplained after controlling for covariates.

[5] However, while these estimates are similar in magnitude and variance, the p value from a test of their equality is 0.52.

[6] This implies that the use of SynthDiD as a preemptive correction may be analogous to the use of heteroskedasticity-robust standard errors. The estimator performs nearly the same as the standard methodology when the correction is not needed, but removes bias when the correction is merited.

Table 2
Mean estimated $\hat{\tau}$.

	Dependent variable:			
	Simulated per-capita armed robbery rate			
	SynthDiD	TWFE	CS	Partially-pooled SCM
$\hat{\tau}_{additiveFE}$	−20.011	−19.984	−20.013	−19.990
	(0.223)	(0.440)	(0.903)	(1.429)
$\hat{\tau}_{Factor}$	−19.926	−19.962	−19.924	−19.916
	(0.179)	(0.286)	(0.329)	(3.015)
Number of simulations			1000	

Note: *p<0.1; **p<0.05; ***p<0.01
Standard errors are in parentheses.

that of a latent factor model, provides evidence that SynthDiD can consistently provide more precise estimates.[7]

6. Conclusion

In this note, I have formalized the functional form of the SynthDiD estimator in settings with staggered treatment timing and have demonstrated its practical use. While, the brief empirical demonstration is in no way meant to be a robust replication of Abrams (2012a,b), it does demonstrate that different aggregated average treatment effect estimates can be obtained using this new and arguably more precise estimator. The simulation exercise demonstrates that in such settings SynthDiD performs as well as traditional DiD methods. However, in more realistic settings in which the underlying data generating process does not include simple additive fixed effects SynthDiD may be able to better estimate the average treatment effect.

SynthDiD may be an appropriate estimator in settings in which the underlying data generating process follows that of a latent factor model or additive fixed effects. However, due to the added computational burden of SynthDiD estimation, this estimator's usage may be inappropriate in big data settings in which the underlying data generating process follows an additive fixed effects model since both methods would be able to consistently estimate the average treatment effect, but with TWFE being much less computationally costly. This brief note provides a base for the implementation of SynthDiD in staggered treatment adoption settings, and hopefully a foundation for this estimator's further development.

Data availability

Data will be made available on request.

Acknowledgments

I would like to express my thanks to Adam Nowak, Feng Yao, Bryan McCannon, and Pedro Sant'Anna, as well as the referee and editor, for their helpful comments and suggestions.

[7] The p value from a test of equality of the SynthDiD and TWFE estimates from this simulation exercise is 0.46.

References

Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. J. Amer. Statist. Assoc. 105 (490), 493–505.
Abadie, A., Gardeazabal, J., 2003. The economic costs of conflict: A case study of the basque country. Amer. Econ. Rev. 93 (1), 113–132. http://dx.doi.org/10.1257/000282803321455188.
Abrams, D.S., 2012a. Estimating the deterrent effect of incarceration using sentencing enhancements. Am. Econ. J.: Appl. Econ. 4 (4), 32–56. http://dx.doi.org/10.1257/app.4.4.32.
Abrams, D.S., 2012b. Replication data for: Estimating the deterrent effect of incarceration using sentencing enhancements. openICPSR. Retrieved January 18, 2022, from https://www.openicpsr.org/openicpsr/project/113838/version/V1/view.
Arkhangelsky, D., Athey, S., Hirshberg, D.A., Imbens, G.W., Wager, S., 2021. Synthetic difference-in-differences. Amer. Econ. Rev. 111 (12), 4088–4118. http://dx.doi.org/10.1257/aer.20190159.
Baker, A., Larcker, D.F., Wang, C.C., 2021. How much should we trust staggered difference-in-differences estimates? SSRN Electron. J. http://dx.doi.org/10.2139/ssrn.3794018.
Ben-Michael, E., Feller, A., Rothstein, J., 2021. Synthetic Controls with Staggered Adoption (No. W28886). National Bureau of Economic Research.
Callaway, B., Sant'Anna, P.H.C., 2021. Difference-in-differences with multiple time periods. J. Econometrics 225 (2), 200–230. http://dx.doi.org/10.1016/j.jeconom.2020.12.001.
De Chaisemartin, C., D'Haultfœuille, X., 2020. Difference-in-differences estimators of intertemporal treatment effects. Available at SSRN 3731856.
Erickson, T., Whited, T.M., 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. Econom. Theory 18 (3), 776–799.
Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. J. Econometrics 225 (2), 254–277. http://dx.doi.org/10.1016/j.jeconom.2021.03.014.
He, G., Wang, S., 2017. Do college graduates serving as village officials help rural China? Am. Econ. J.: Appl. Econ. 9 (4), 186–215. http://dx.doi.org/10.1257/a20160079.
Kahn, J., 2015. Influence Functions for Fun and Profit. Ross School of Business, University of Michigan, Available from http://j-kahn.com/files/influencefunctions.pdf (version: July 10, 2015).
Roth, J., 2022. Pre-test with caution: Event-study estimates after testing for parallel trends. Am. Econ. Rev.: Insights 4 (3), 305–322. http://dx.doi.org/10.1257/aeri.20210236.
Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J. Econometrics 225 (2), 175–199. http://dx.doi.org/10.1016/j.jeconom.2020.09.006.