

Design-Based Analysis in Difference-In-Differences Settings with Staggered Adoption, Journal of Econometrics, 2022

Wenzhi Wang *

August 20, 2024

Athey and Imbens (2022)

1. Introduction

In this paper we study estimation of and inference for average treatment effects in a setting with panel data. We focus on the setting where units adopt the policy or treatment of interest at a particular point in time, and then remain exposed to this treatment at all times afterwards. The adoption date at which units are first exposed to the policy may, but need not, vary by unit. We refer to this as a *staggered adoption design (SAD)*, such designs are sometimes also referred to as event study designs.

This paper takes a *design-based* perspective where the properties of the estimators arises from the stochastic nature of the treatment assignment, rather than a *sampling-based* or *model-based* perspective where these properties arise from the random sampling of units from a large population in combination with assumptions on this population distribution. This perspective is particularly attractive in the settings when the sample comprises the entire population, e.g., all states of the US, or all countries of the world so that nondegenerate sampling properties would require postulating an imaginary super-population.

In this design setting our critical assumptions involve restrictions on the assignment process as well as exclusion restrictions, but in contrast to other work in this area they do not include functional form assumptions. Commonly made common trend assumptions follow from some of our assumptions, but are not the starting point.

We set up the problem with the adoption date, rather than the actual exposure to the intervention, as the basic treatment indexing the potential outcomes. We consider assumptions under which this

*This note is written in my pre-doc period at the University of Chicago Booth School of Business.

discrete multi-valued treatment (the adoption date) can be reduced to a binary one, defined as the indicator whether or not the treatment has already been adopted. We then investigate the properties of the standard DID estimator under assumptions about the assignment of the adoption date and under various exclusion restrictions. We show that under a random adoption date assumption, the standard DID estimator can be interpreted as the weighted average of average causal effects of changes in the adoption date. We also consider design-based inference for this estimand.

2. Set up

Using the potential outcome framework for causal inference, we consider a setting with a finite population of N units. Each of these N units are characterized by a set of potential outcomes in T periods for $T + 1$ treatment levels, $Y_{it}(a)$. Here $i \in \{1, \dots, N\}$ indexes the units, $t \in \mathbb{T} = \{1, \dots, T\}$ indexes the time periods, and the argument of the potential outcome function $Y_{it}(\cdot)$ is the adoption date $a \in \mathbb{A}$, where $\mathbb{A} = \mathbb{T} \cup \{\infty\} = 1, \dots, T, \infty$. This argument, a , which indexes the discrete treatment, is the date that the binary policy was first adopted by a unit. Units can adopt the policy at any of the time periods $1, \dots, T$, or not adopt the policy at all during the period of observation, in which case we code the adoption date as ∞ . Once a unit adopts the treatment, it remains exposed to the treatment for all periods afterwards.

This set up differs from that in most of the DiD literature where the binary indicator whether a unit is exposed to the treatment in the current period indexes the potential outcomes. The notion of focusing on a full treatment path rather than a binary treatment is also related to the dynamic treatment effect literature. We observe for each unit in the population the adoption date $A_i \in \mathbb{A}$ and the sequence of T realized outcomes, Y_{it} , for $t \in \mathbb{T}$, where the realized outcome for unit i in period t equals

$$Y_{it} = Y_{it}(A_i). \quad (1)$$

We may also observe pre-treatment characteristics, denoted by the K -component vector X_i , although for most of the discussion we abstract from their presence. Let \mathbf{Y} , \mathbf{A} , and \mathbf{X} denote the $N \times T$, $N \times 1$ and $N \times K$ matrices with typical elements Y_{it} , A_i , and X_{itk} , respectively.

Our design-based analysis views the potential outcomes $Y_{it}(a)$ as deterministic, and only the adoption dates A_i , as well as functions thereof such as the realized outcomes, as stochastic. Distributions of estimators will be fully determined by the adoption date distribution, with the number of units N and the number of time periods T fixed, unless explicitly stated otherwise.

In many cases the units themselves are clusters of units of a lower level of aggregation. For example, the units may be states, and the outcomes could be averages of outcomes for individuals in that state, possibly of samples drawn from subpopulations from these states. In such cases N and T may be as small as 2, although in many of the cases we consider N will be at least moderately large. This distinction between cases where Y_{it} is itself an average over basic units or not, affects some, but not all, of the formal statistical analyses. It may make some of the assumptions more plausible, and it may affect the inference, especially if individual level outcomes and covariates are available.

Define $W : \mathbb{A} \times \mathbb{T} \mapsto \{0, 1\}$, with $W(a, t) = \mathbb{I}(a \leq t)$ to be the binary indicator for the adoption date a preceding t , and define W_{it} to be the indicator for the policy having been adopted by unit i prior to, or at, time t , so that

$$W_{it} = W(A_i, t) = \mathbb{I}(A_i \leq t).$$

The $N \times T$ matrix \mathbf{W} with typical element W_{it} has the form:

$$\mathbf{W}_{N \times T} = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & T & \text{(time period)} \\ 0 & 0 & 0 & 0 & \dots & 0 & \text{(never adopter)} \\ 0 & 0 & 0 & 0 & \dots & 1 & \text{(late adopter)} \\ 0 & 0 & 0 & 0 & \dots & 1 & \\ 0 & 0 & 1 & 1 & \dots & 1 & \\ 0 & 0 & 1 & 1 & \dots & 1 & \text{(medium adopter)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 1 & 1 & 1 & \dots & 1 & \text{(early adopter)} \end{pmatrix}$$

Let $N_a = \sum_{i=1}^N \mathbb{I}(A_i = a)$ be the number of units in the sample with adoption date a , define $\pi_a = N_a/N$, for $a \in \mathbb{A}$, as the fraction of units with adoption date equal to a , and define $\Pi_t = \sum_{s=1}^t \pi_s$, for $t \in \mathbb{T}$, as the fraction of units with an adoption date on or prior to t .

Also define $\bar{Y}_t(a)$ to be the population average of the potential outcome in period t for adoption date a :

$$\bar{Y}_t(a) \equiv \frac{1}{N} \sum_{i=1}^N Y_{it}(a), \quad \text{for } t \in \mathbb{T}, a \in \mathbb{A}.$$

Define the unit and average causal effects of adoption date a' relative to a , on the outcome in period t , as

$$\tau_{it,aa'} \equiv Y_{it}(a') - Y_{it}(a), \quad \tau_{t,aa'} \equiv \frac{1}{N} \sum_{i=1}^N \{Y_{it}(a') - Y_{it}(a)\} = \bar{Y}_t(a') - \bar{Y}_t(a).$$

The average causal effects $\tau_{t,aa'}$ form components of many of the estimands we consider later. A particularly interesting average effect is

$$\tau_{t,\infty 1} = \frac{1}{N} \sum_{i=1}^N (Y_{it}(1) - Y_{it}(\infty)),$$

the average effect in period t of switching the entire population from never adopting the policy ($a = \infty$), to adopting the policy in the first period ($a = 1$). Formally, there is nothing special about the particular average effect $\tau_{t,\infty 1}$ relative to any other $\tau_{t,aa'}$, but $\tau_{t,\infty 1}$ will be useful as a benchmark. Part of the reason is that for all t and i the comparison $Y_{it}(1) - Y_{it}(\infty)$ is between potential outcomes for adoption prior to or at time t (namely adoption date $a = 1$) and potential outcomes for adoption later than t (namely, never adopting, $a = \infty$). In contrast, any other average effect $\tau_{t,aa'}$ will for some t involve comparing potential outcomes neither of which correspond to having adopted the treatment yet, or comparing potential outcomes both of which correspond to having adopted the treatment already.

3. Assumptions

We consider three sets of assumptions. The first set, containing only a single assumption, is about the *design*, that is, the assignment of the treatment, here the adoption date, conditional on the potential outcomes and possibly pretreatment variables. We refer to this as a design assumption because it can be guaranteed by design of the study. The second set of assumption is about the *potential outcomes*, and rules out the presence of certain treatment effects. These exclusion restrictions are substantive assumptions, and they cannot be guaranteed by design. The third set of assumptions consists of four *auxiliary assumptions*, two about heterogeneity of certain causal effects, one about sampling from a large population, and one about an outcome model in a large population. The nature of these three sets of assumptions, and their plausibility, is very different, and it is in our view useful to carefully distinguish between them.

3.1. The Design Assumption

The first assumption is about the assignment process for the adoption date A_i . Our starting point is to assume that the adoption date is completely random:

Assumption 1 (Random Adoption Date). For some set of positive intergers N_a , for $a \in \mathbb{A}$,

$$\mathbb{P}(\mathbf{A} = \mathbf{a}) = \left(\frac{N!}{\prod_{a \in \mathbb{A}} N_a!} \right)^{-1},$$

for all N -vectors \mathbf{a} such that for all $a \in \mathbb{A}$, $\sum_{i=1}^N \mathbb{I}_{a_i=a} = N_a$.

This assumption is obviously very strong. However, without additional assumptions, this assumption has no testable implications in a setting with exchangeable units, as stated formally in the following Lemma.

Lemma 1 (No Testable Restrictions). Suppose all units are exchangeable. Then Assumption 1 has no testable implications for the joint distribution of (\mathbf{Y}, \mathbf{A}) .

Hence, if we wish to relax the assumptions, we need to bring in additional information. Such additional information can come in the form of pretreatment variables, that is, variables that are known not to be affected by the treatment. In that case we can relax the assumption by requiring only that the adoption date is completely random within subpopulations with the same values for the pre-treatment variables. Additional information can also come in the form of restrictions on the potential outcomes or the treatment effects.

Under Assumption 1 the marginal distribution of the adoption dates is fixed, and so also the fraction π_a is fixed in the repeated sampling thought experiment. This part of the set up is similar in spirit to

fixing the number of treated units in the sample in a completely randomized experiment. It is convenient for obtaining finite sample results. *Note that it implies that the adoption dates for units i and j are not independent.*

An important role in our analysis is played by what we label the *adjusted treatment*, adjusted for unit and time period averages:

$$\dot{W}_{it} \equiv W_{it} - \bar{W}_{\cdot t} - \bar{W}_{i \cdot} + \bar{W},$$

where $\bar{W}_{\cdot t} \equiv \sum_{i=1}^N W_{it}/N$, $\bar{W}_{i \cdot} \equiv \sum_{t=1}^T W_{it}/T$, and $\bar{W} \equiv \sum_{i=1}^N \sum_{t=1}^T W_{it}/(NT)$ are averages over units, time periods, and both, respectively. We can also write the adjusted treatment indicator as

$$\dot{W}_{it} = g(t, A_i),$$

where

$$g(t, a) \equiv \left(\mathbb{I}(a \leq t) - \sum_{s \leq t} \pi_s \right) + \frac{1}{T} \left(a \mathbb{I}(a \leq T) - \sum_{s=1}^T s \pi_s \right) + \frac{T+1}{T} (\mathbb{I}(a = \infty) - \pi_\infty)$$

where, with some minor abuse of notation, we adopt the convention that $a \mathbb{I}(a \leq T)$ is zero if $a = \infty$. Under Assumption 1, which fixes the marginal distribution of A_i , the sum $\sum_{i,t} \dot{W}_{it}^2$ is non-stochastic, even though the adjusted treatment \dot{W}_{it} and its squares are stochastic. The fact that this sum is non-stochastic enables us to derive exact finite sample results for the standard DID estimator.

3.2. Exclusion Restrictions

The next two assumptions concern the potential outcomes. Their formulation does not involve the assignment mechanism, that is, the distribution of the adoption date. In that sense they are unlike the strict and weak exogeneity or no-feedback assumptions. In essence these are exclusion restrictions, assuming that particular causal effects are absent. Collectively these two assumptions imply that we can think of the treatment as a binary one, the only relevant component of the adoption date being whether a unit is exposed to the treatment at the time we measure the outcome.

The first of the two assumptions, and likely the more plausible of the two in practice, rules out effects of future adoption dates on current outcomes. More precisely, it assumes that if the policy has not been adopted yet, the exact future date of the adoption has no causal effect on potential outcomes for the current period.

Assumption 2 (No Anticipation). For all units i , all time periods t , and for all adoption dates a , such that $a > t$,

$$Y_{it}(a) = Y_{it}(\infty).$$

We can also write this assumption as requiring that for all triples (i, t, a) ,

$$Y_{it}(a) = \mathbb{I}(a \leq t)Y_{it}(a) + \mathbb{I}(a > t)Y_{it}(\infty), \quad \text{or } \mathbb{I}(a > t)(Y_{it}(a) - Y_{it}(\infty)) = 0.$$

This last representation shows most clearly how the assumption rules out certain causal effects. Note that this assumption does not involve the adoption date, and so does not restrict the distribution of the adoption dates. Violations of this assumption may arise if the policy is anticipated prior to its implementation.

The next assumption is arguably much stronger. It asserts that for potential outcomes in period t it does not matter how long the unit has been exposed to the treatment, only whether the unit is exposed at time t .

Assumption 3 (Invariance to History). For all units i , all time periods t , and for all adoption dates a , such that $a \leq t$,

$$Y_{it}(a) = Y_{it}(1).$$

This assumption can also be written as

$$Y_{it}(a) = \mathbb{I}(a \leq t)Y_{it}(1) + \mathbb{I}(a > t)Y_{it}(a), \quad \text{or } \mathbb{I}(a \leq t) (Y_{it}(a) - Y_{it}(1)) = 0$$

with again the last version of the assumption illustrating the exclusion restriction in this assumption. Again, the assumption does not rule out any correlation between the potential outcomes and the adoption date, only that there is no causal effect of an early adoption versus a later adoption on the outcome in period t , as long as adoption occurred before or on period t .

If both the exclusion restrictions, that is, both Assumptions 2 and 3 hold, then the potential outcome $Y_{it}(a)$ can be indexed by the binary indicator $W(a, t) = \mathbb{I}(a \leq t)$.

Lemma 2 (Binary Treatment). Suppose Assumptions 2 and 3 hold.

(1) For all units i , all time periods t , and for all adoption dates $a \geq a'$,

$$Y_{it}(a') - Y_{it}(a) = \mathbb{I}(a' \leq t < a) \times (Y_{it}(1) - Y_{it}(\infty)),$$

so that

$$Y_{it}(a) = Y_{it}(\infty) + \mathbb{I}(a \leq t) \times (Y_{it}(1) - Y_{it}(\infty)) = \begin{cases} Y_{it}(\infty) & \text{if } a \leq t \\ Y_{it}(1) & \text{otherwise} \end{cases}.$$

(2) For all time periods t , and adoption dates $a > a'$,

$$\tau_{t,aa'} = \tau_{t,\infty 1} \mathbb{I}(a' \leq t < a) = \begin{cases} \tau_{t,\infty 1} & \text{if } a' \leq t < a, \\ 0 & \text{otherwise.} \end{cases}.$$

If these two assumptions hold, we can therefore simplify the notation for the potential outcomes and focus on the pair of potential outcomes $Y_{it}(1)$ and $Y_{it}(\infty)$.

Note that these two assumptions are substantive, and because they only involve the potential outcomes and not the adoption date, they cannot be guaranteed by design. This is in contrast to Assumption 1, which can be guaranteed by randomization of the adoption date. It is also important to note that in many empirical studies Assumptions 2 and 3 are made, often implicitly by writing a model for realized outcome Y_{it} that depends solely on the contemporaneous treatment exposure W_{it} , and not the actual adoption date A_i or treatment exposure $W_{it'}$ in other periods t' . In the current discussion we want to be explicit about the fact that this restriction is an assumption, and it does not automatically hold. Note that the assumption does not restrict the time series dependence between the potential outcomes.

It is trivial to see that without additional information, the exclusion restrictions in Assumptions 2 and 3 have no testable implications because they impose restrictions on pairs of potential outcomes that can never be observed together. However, in combination with random assignment, the two exclusion restrictions, Assumptions 2 and 3, have testable implications as long as $T \geq 2$ and there is some variation in the adoption date.

Lemma 3 (Testable Restrictions from the Exclusion Restrictions). (1) Assumptions 2 and 3 jointly have no testable implications for the joint distribution of (\mathbf{Y}, \mathbf{W}) .
 (2) Suppose $T \geq 2$ and $\pi_2, \pi_\infty > 0$. Then the combination of the random adoption date and the exclusion restrictions, Assumptions 1-3, impose testable restrictions on the joint distribution of (\mathbf{Y}, \mathbf{W}) .

This implies that if we maintain, say, the no-anticipation assumption, we can relax the random adoption date assumption. For example, we can allow the probability of adoption at time t to depend on the outcomes prior to period t .

3.3. Auxiliary Assumptions

In this section we consider four auxiliary assumptions that are convenient for some analyses, and in particular can have implications for the variance of specific estimators, but that are not essential in many cases. These assumptions are often made in empirical analyses without researchers explicitly discussing them

The first of these assumptions assumes that the effect of adoption date a' , relative to adoption date a , on the outcome in period t , is the same for all units.

Assumption 4 (Constant Treatment Effects over Units). For all units i, j and for all time periods t and all adoption dates a and a' ,

$$Y_{it}(a') - Y_{it}(a) = Y_{jt}(a') - Y_{jt}(a).$$

The second assumption restricts the heterogeneity of the treatment effects over time.

Assumption 5 (Constant Treatment Effect over Time). For all units i and all time periods t and t'

$$Y_{it}(1) - Y_{it}(\infty) = Y_{it'}(1) - Y_{it'}(\infty).$$

We only restrict the time variation for comparisons of the adoption dates 1 and ∞ because we typically use this assumption in combination with Assumptions 2 and 3. In that case we obtain a constant binary treatment effect set up, as summarized in the following Lemma.

Lemma 4 (Binary Treatment and Constant Treatment Effects). Suppose Assumptions 2-5 hold. Then for all t and $a' < a$

$$Y_{it}(a') - Y_{it}(a) = \mathbb{I}(a' \leq t < a) \times \tau_{1\infty}.$$

The next assumption allows us to view the potential outcomes by postulating a large population from which the sample is drawn.

Assumption 6 (Random Sampling). The sample can be viewed as a random sampling from an infinitely large population, with joint distribution for $(A_i, Y_{it}(a), a \in \mathbb{A}, t \in \mathbb{T})$ denoted by $f(a, y_1(1), \dots, y_T(\infty))$.

Under the random sampling assumption we can put additional structure on average potential outcomes.

Assumption 7 (Additivity). $\mathbb{E}[Y_{it}(\infty)] = \alpha_i + \beta_t$.

4. Difference-in-Differences Estimators: Interpretation and Inference

In the simplest setting with N units and T time periods, without additional covariates, the realized outcome in period t for unit i is modelled as

$$Y_{it} = \alpha_i + \beta_t + \tau W_{it} + \varepsilon_{it}. \quad (2)$$

In this model there are unit effects α_i and time effects β_t , but both are additive with interactions between them ruled out. The effect of the treatment is implicitly assumed to be additive and constant across units and time periods.

We interpret the DiD estimand under the randomized adoption date assumption, leading to a different setting from other econometrics papers.

4.1. Difference-in-Differences Estimators

References

Athey, Susan and Guido W. Imbens (2022) "Design-Based Analysis in Difference-in-Differences Settings with Staggered Adoption," *Journal of Econometrics*, 226 (1), 62–79, [10.1016/j.jeconom.2020.10.012](https://doi.org/10.1016/j.jeconom.2020.10.012).