Learning and Identification in DiD designs

Manuel Soffici

Department of Statistics, Ludwig-Maximilians-Universität München,
80799 Munich, Germany, sofficimanuel@gmail.com

16th March 2025

Abstract

Difference-in-Differences is a popular design for estimating treatment effects, especially in policy evaluation. Its central identifying assumption is the Parallel Trends (PT) condition: the trend in mean untreated outcomes is independent of the observed treatment status. In observational settings, treatment is often a dynamic choice determined by rational agents, such as policy-makers or individual actors. This paper is based on the work of Marx et al., 2024, who examine the relationship between PT and economic models of dynamic choice. They make explicit the constraints PT puts on agent behavior and study when dynamic behavior violates the assumption. My paper focuses on a subset of dynamic selection scenarios, namely mechanisms of learning about potential outcomes. This work revisits the theoretical framework defined by Marx et al., 2024, interprets the findings intuitively, and demonstrates the main ideas and results with a novel simulation.

Keywords: Difference-in-differences, parallel trends, dynamic choice models, treatment effects, causal inference, selection mechanism.

1. Introduction

Difference-in-Differences (DiD) is a standard non-experimental research design for causal inference on panel or repeated cross-section data. The canonical DiD setting is one where outcomes are observed for subjects belonging to one of two groups (treatment or control) in two distinct time periods (pre-treatment and post-treatment). In the first period no one is treated, while in the second period only units in the treatment group are. The average outcome gain over time, or *trend*, in the control group is subtracted from the trend in the treatment group. If subjects were randomly assigned to the two groups with perfect compliance, this double differencing would remove biases in post-treatment comparisons between the treatment and control group arising from systematic differences between them. The DiD estimator would then be an unbiased estimator of the *Average*

Treatment effect on the Treated (ATT), a simple and popular causal estimand.

It is widely known that social science research mostly relies on non-experimental or quasi-experimental studies, therefore the random assignment assumption must find an appropriate theoretical surrogate in its place. The validity of a formally untestable hypothesis, Parallel Trends (PT), is arguably sufficient to point-identify the causal parameter of interest in the vast majority of studies. In canonical sharp designs, ie. designs with a pre-treatment period during which no unit is treated, PT simply maintains that, absent treatment, treated and untreated subjects would have followed a common trend, as shown in Figure 1. As a consequence, explicit assumptions on treatment selection mechanisms are not necessary for identification, and this is one of the reasons for the wide popularity of simple DiD designs in social sciences.

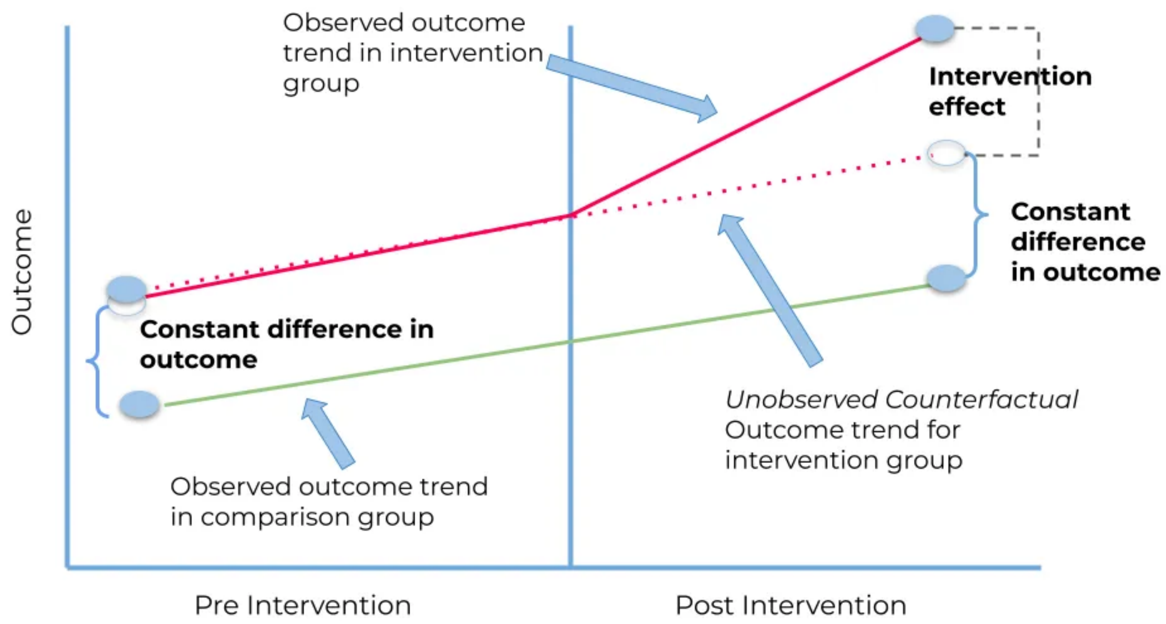

Figure 1: Parallel Trends in sharp DiD designs. (Source)

1.1 Literature review.

Over the past few decades, econometricians have increasingly recognized the challenges of causal identification in complex DiD designs, especially when extending to multiple time periods or fuzzy designs, ie. designs where units may be treated in the first period and the notion of group allocation is not (as) meaningful. This emergent awareness is attested by the recent publication of "a dizzying array of new methodological papers on DiD and related designs", reviewed by Roth et al., 2023, de Chaisemartin et al., 2024, and Arkhangelsky & Imbens, 2024. A historical overview of the literature on canonical DiD designs can instead be found in Section 2 of Lechner et al., 2011.

Recent works have critiqued careless reliance on poorly justified PT assumptions, and sought to provide alternative pathways to (partial) identification for when the PT condition is not met. For example, Manski & Pepper, 2018, Ban & Kédagni, 2023, and Rambachan & Roth, 2023 propose partial identification under weaker assumptions. Abadie, 2005 considers identification under covariate-conditional parallel trends for when treatment effect is modified by covariates. In the presence of pre-trends, Dobkin et al., 2018 consider extrapolations of pre-trends as an alternative to PT, while Freyaldenhoven et al., 2019 propose an alternative identification approach based on covariates relating to the examined policy only through a pre-trend confounder. Complementary to this work, our discussion of identification focuses on alternatives that do not require any additional data (such as pre-trends, proxies, or instrumental variables), similarly to the doubly-robust approach to identification in panel data models of Arkhangelsky & Imbens, 2022.

A particular strand of this literature focuses on developing alternative estimation methods specifically by imposing explicit selection and/or outcome models. Prominent works in the field have been published by de Chaisemartin & D'HaultfŒuille, 2017, Verdier, 2020, and Ghanem et al., 2024. Among these, the latter is the closer in spirit to the work of Marx et al., 2024. Ghanem et al., 2024 now provide formal guidance for practitioners on how to evaluate the PT assumption. Their main result is that PT can be compatible with restricted selection on time-varying unobservables and they encourage researchers to justify the compatibility of selection mechanisms and PT with contextual or domain knowledge.

Our contribution. Since in these settings subjects may make new treatment choices after observing post-treatment outcomes, dynamic selection mechanisms may arise. Marx et al., 2024 complement this literature by exploring the relationship between causal identification and structural models of dynamic choice. Further insights come from Fudenberg & Levine, 2022, who emphasize how learning by decision-makers complicates the interpretation of natural experiments, and from Vytlacil, 2002, who establish equivalences across various choice models, thus clarifying the implicit assumptions underlying selection processes.

Moreover, Marx et al., 2024 take considerable inspiration from the literature that explicitly models dynamic decisions with incentives, learning, and/or option values. See for example Taber, 2001, Heckman & Navarro, 2007, and Abbring, 2010. Because models from this literature reflect the behavior of economic optimizing agents, they embed causal parameters in themselves. Thus, causal relations are provided by construction through model specifications that can be used to address counterfactuals and policy impacts. DiD research focuses instead on identifying parameters that are given a causal interpretation under some restrictive assumptions like PT, without the need for full specification of a behavioral model. This is an undeniable strength of DiD, but researchers should ensure their identifying assumptions are not compromised by dynamic selection mechanisms.

Finally, my work revisits the theoretical framework defined by Marx et al., 2024. In particular, it provides an intuitive exposition of their theoretical insights and supports these results through simulation-based analyses. Our simulations explicitly demonstrate scenarios where learning about potential outcomes can lead to systematic violations of the PT assumption, thereby biasing traditional DiD estimates. This empirical demonstration reinforces the practical significance of carefully understanding the types of learning mechanisms operating within DiD studies. It provides a clear blueprint for researchers to assess the validity of PT assumptions within empirical applications, particularly those where decision-making is influenced by evolving preferences or information acquisition over time.

Our contribution is especially pertinent to empirical contexts where dynamic selection driven by preferences or information updates is likely to play a prominent role. By illuminating how specific learning mechanisms can invalidate PT, we equip researchers with theoretical and empirical tools to diagnose and address potential biases in their causal analyses.

Let us outline the structure of this paper. First, we establish our analytical framework, defining essential notation, formally specifying our dynamic choice model, and illustrating key concepts with a clear and intuitive example. Subsequently, we examine three distinct learning scenarios to analyze how the acquisition of new information influences the validity of the PT assumption. We then provide a concrete empirical demonstration through a simulation study based on a job training program. Finally, we synthesize our theoretical and empirical findings, discuss their implications for applied DiD research, and outline practical guidelines for researchers aiming to validate PT assumptions under dynamic selection settings.

2. Setup

Consider a job training program, analogous to the one presented by Ashenfelter & Card, 1985, as our running case study. Simultaneously, let us introduce a formalized model specification and notation to systematically define the framework.

We observe a randomly drawn sample of N workers, indexed by i, who are eligible for the job training program. Participation in the program at time t is denoted by D_{it}, where $D_{it} = 1$ indicates that worker i enrolls in the program, while $D_{it} = 0$ represents non-participation. Importantly, enrollment is flexible and permitted in both periods, ie. any $(d_{i0}, d_{i1}) = \mathbf{d}_i$ is possible, hence there is *no strictly defined pre-treatment period*. This structure suggests a program of variable duration or a recurring intervention offered across two distinct sessions.

The primary outcome variable Y_{it}, is measured at two time points, $t = 0$ and $t = 1$. In our case they represent wages or salaries, which are continuous quantities. Each observed outcome $Y_{it} = D_{it}Y_{it}(1) + (1 - D_{it})Y_{it}(0)$ corresponds to one of two potential (or counterfactual) outcomes $Y_{it}(D_{it})$, depending on the individual's treatment status.

A fundamental assumption in this setup is that an individual's outcome at time t depends solely on their treatment status at that time: $Y_{it}(\mathbf{d}_1, \ldots, \mathbf{d}_N) = Y_{it}(d_{it})$. Two important implications follow. First, there are *no spillover effects or network dependencies*: the outcomes of individuals are assumed to be independent of each other. Combined with the assumption of random sampling and a well-defined treatment, this guarantees that the *Stable Unit Treatment Value Assumption* (SUTVA) holds. Consequently, we can analyze population-level variables without employing individual-specific subscripts. Moreover, there is no *no dependence on treatment history*: the realized outcome at time t depends exclusively on the contemporaneous treatment status D_{it}, rather than the entire treatment history. This assumption effectively rules out dynamic treatment effects and anticipatory responses to future treatment status.

While these restrictions are nontrivial, they are standard in the literature and will not be further elaborated upon here.

2.1 Model of dynamic choice.

A key challenge in modeling treatment decisions is accounting for the *dynamic choices* made by rational agents. In this section, we develop a framework that captures these choices by incorporating utility maximization, treatment costs, and information updating.

The model is structured around several components:

- *Potential outcomes.* The primary variable of interest is the potential outcome $Y_t(d_t)$, which represents the observed outcome under treatment status d_t. If wages serve as the outcome, they can be directly compared with the costs of program enrollment. However, for binary outcomes like employment status, it is necessary to map them into a real-valued function allowing comparability with treatment costs.

- *Treatment costs.* Each treatment decision is associated with a history-dependent cost, denoted as $K_t(\mathbf{d}_t)$. Note that $K_0(D_0)$ denotes the cost of treatment D_0 for $t = 0$, while $K_1(d_0, D_1)$ denotes the cost of treatment D_1 for $t = 1$ if $D_0 = d_0$.

- *Per-period utility.* The utility derived from treatment at time t is defined as the difference between the potential outcome and the cost of treatment:

$$V_t(d_t) = Y_t(d_t) - K(d_t). \tag{1}$$

- *Expected future utility.* Forward-looking behavior implies that individuals anticipate future benefits when making treatment decisions. To formalize this, we introduce the expected future utility of the optimal treatment decision at $t = 1$, denoted as $W_1(d_0)$. The best decision at $t = 1$ is determined based on the updated information set $U_1(d_0)$, which captures all relevant knowledge accumulated by the start of period 1. We define it as the random variable:

$$W_1(d_0) \equiv \max_{d_1 \in \{0,1\}} E[V_1(d_1)|U_1(d_0)] \tag{2}$$

- *Discounting of future gains.* Since decisions are made dynamically, future utility must be translated into present value terms. We introduce a discount factor $\beta \in (0, 1)$, where a value of $\beta = 0.5$ implies that a future gain of 100€ is equivalent to a present gain of 50€.

- *Information updating and learning.* At time $t = 0$, the agent's available information is denoted as U_0, which includes the discount factor β and treatment costs $K(d_t)$. However, the agent does not have complete knowledge of potential outcomes. Over time, information updates based on observed realizations $Y_0 = Y_0(d_0)$, forming the updated information set $U_1(d_0)$. Information does not decrease: $U_0 \subseteq U_1(d_0)$.

The distinction between forward-looking behavior and learning is crucial. The variables $W_1(d_0)$ and β capture *forward-looking decision-making*, while the information sets model *learning from past outcomes*, which in turn allows agents to experiment with treatment decisions and adjust their choices.

Having established the model components, we now define the decision-making process formally. The Dynamic Utility Maximization (DUM) framework assumes that agents make treatment choices by maximizing the sum of expected discounted utilities. The decision process follows an inductive structure, meaning that the choice rule for the second treatment decision D_1 is nested within the rule for the first treatment decision D_0.

Dynamic Utility Maximization (DUM). Given initial information U_0 at the start of period 0 and accrued information $U_1(d_0)$ at the start of period 1, treatment decisions maximize the sum of expected discounted utility:

$$\begin{cases} D_1(d_0) = 1\{E[V_1(1) - V_1(0)|U_1(d_0)] \geq 0\} \\ \quad D_0 = 1\{E[V_0(1) - V_0(0) + \beta(W_1(1) - W_1(0))|U_0] \geq 0\}, \end{cases}$$

where $1\{\cdot\}$ denotes the indicator function and $U_1(d_0) \equiv \{U_0, Y_0(d_0)\}$.

The decision rule for treatment at $t = 1$ is relatively straightforward. Given the updated information $U_1(d_0)$, the individual chooses D_1 by comparing the expected utilities associated with the two possible treatment states:

$$D_1(d_0) = 1\{E[V_1(1) - V_1(0)|U_1(d_0)] \geq 0\}. \tag{3}$$

This equation states that if the expected net utility from treatment at $t = 1$ is positive, the agent opts for treatment ($D_1 = 1$); otherwise, they do not ($D_1 = 0$).

The decision rule for D_0 is slightly more complex since it must account for both immediate utility and the expected future utility of treatment decisions at $t = 1$. Given the initial information set U_0, the agent chooses D_0 according to:

$$D_0 = 1\{E[V_0(1) - V_0(0) + \beta(W_1(1) - W_1(0))|U_0] \geq 0\}. \tag{4}$$

The first term, $V_0(1) - V_0(0)$, captures the immediate utility gain from treatment at $t = 0$. The second term, $W_1(1) - W_1(0)$, represents the difference in future expected utilities under treatment and non-treatment at $t = 0$. The discount factor β scales the impact of future gains on present decision-making.

Since the model is inductive, it assumes that the agent knows the optimal D_1 in the first period, while in reality future decisions may still be uncertain. This assumption allows us to incorporate dynamic optimization within a tractable framework.

2.2 Model illustration.

The model formulation highlights that learning dynamics can encourage experimentation with treatment. If agents face uncertainty about potential gains, they may opt (not) to enroll in training at $t = 0$ as a way of gathering information that informs their future decisions. Understanding these mechanisms is essential when evaluating treatment effects, as experimentation motivated by learning may influence selection into treatment, potentially challenging the Parallel Trends assumption.

To provide intuition for the dynamic utility maximization framework, we consider a subject-level toy example. In this setting, expectations regarding potential outcomes are interpreted as subjective beliefs. Treatment costs are not history-dependent, so we ignore the time subscripts. In subsequent sections, we will deal with population-level variables.

Consider a worker, Bob, who must decide whether to enroll in a job training program. His initial expectations about his salary outcomes are as follows:

- he *expects* a wage of $300 = E[Y_0(1)|U_0]$ after enrolling in $t = 0$, or a wage of $200 = E[Y_0(0)|U_0]$ if he does not enroll in $t = 0$;
- he *knows* enrolling costs $200 = K(1)$, while not enrolling is free, ie. $0 = K(0)$.

Given these expectations, we compute Bob's expected utilities in period $t = 0$:

$$V_0(1) = \mathbb{E}[Y_0(1) \mid U_0] - K(1) = 300 - 200 = 100, \tag{5}$$

$$V_0(0) = \mathbb{E}[Y_0(0) \mid U_0] - K(0) = 200 - 0 = 200. \tag{6}$$

Since $V_0(0) > V_0(1)$, Bob would prefer not to enroll in the program at $t = 0$, if he only considered immediate utility. However, his decision must also take into account the expected future utility associated with each choice. If Bob *values* 50€ today as much as 100€ in the future, we can employ a discount factor $\beta = 0.5$.

To determine the expected utility of the best possible decision at $t = 1$, we assume that Bob already possesses the updated information set $U_1(d_0)$, even though in reality this information only becomes available after observing Y_0. We assume Bob's expectations in $U_1(d_0)$ for wages at $t = 1$ are as follows: he *expects* a salary of $800 = E[Y_1(1)|U_1(d_0)]$ after enrolling in $t = 1$, or instead $300 = E[Y_1(0)|U_1(d_0)]$ if he does not enroll in $t = 1$.

Since enrolling at $t = 1$ yields a considerably higher expected salary than not enrolling, the optimal choice at $t = 1$ is to enroll. Thus, the conditional expectation of $W_1(d_0)$ is given by:

$$W_1(d_0) = 800 - 200 = 600. \tag{7}$$

Bob's decision at $t = 0$ is made, in the DUM framework, according to:

$$\begin{aligned} D_0 &= 1\{E[V_0(1) - V_0(0) + \beta(W_1(1) - W_1(0))|U_0] \geq 0\} \\ &= 1\{100 - 200 + 0.5 * (600 - 600) \geq 0\} \\ &= 1\{-100 \geq 0\} = 0. \end{aligned} \tag{8}$$

Thus, Bob optimally chooses not to enroll in the program at $t = 0$.

Then he observes the realized outcome $Y_0(0)$, which may lead him to update his expectations regarding potential wages at $t = 1$. Consequently, when deciding whether to enroll at $t = 1$, he follows the same decision rule, but this time without needing to discount future gains, since there is no further period to consider.

This example illustrates the fundamental mechanics of rational dynamic choice: agents make treatment decisions by weighing immediate costs and benefits against expected future gains, while learning in the process.

Having established a model of rational dynamic choice, we now return to the central question: under what selection mechanisms does the Parallel Trends (PT) assumption hold? If individuals adjust their treatment choices in response to past outcomes, the assumption that untreated potential outcomes evolve in parallel across groups may no longer be valid. In the following sections, we explore specific selection mechanisms and evaluate how learning about potential outcomes affects the validity of PT. Some learning mechanisms are compatible with PT, while others induce biases that violate it.

3. Learning mechanisms

Before proceeding, it is important to recognize that in studies without a pre-treatment period, violations of the PT assumption can occur at different points in time. Specifically, PT may fail in period $t = 0$, period $t = 1$, or both. This is reflected and formalized in Observation 1 of Marx et al., 2024, who provide a set of conditions equivalent to PT.

Observation 1. The PT condition is equivalent to the following statements

(i) the conditional untreated trend is constant in d_1 for each d_0:

$$E[Y_1(0) - Y_0(0)|d_0, d_1 = 0] = E[Y_1(0) - Y_0(0)|d_0, d_1 = 1] \text{ for any } d_0,$$

ie. *no selection when choosing* D_1,

(ii) the untreated trend conditional on d_0 is constant in d_0:

$$E[Y_1(0) - Y_0(0)|d_0 = 0] = E[Y_1(0) - Y_0(0)|d_0 = 1],$$

ie. *no selection when choosing* D_0,

(iii) the constant trends are all equal, ie. *no trend shift between time 0 and 1.*

These conditions allow us to systematically diagnose *when and where* PT fails in our setting. If PT does not hold, the violation may stem from differences in untreated trends at $t = 0$, at $t = 1$, or both.

An additional clarification is necessary. In this context, when referring to "selection", we generally mean *selection on untreated trends*, as this is what directly affects the validity of PT. The breakdown provided by Observation 1 helps distinguish between different sources of PT violations, enabling us to assess whether trends diverge in the initial period and/or the second period, and in what subgroup.

A key distinction in this setting is between *static selection* (which occurs based on initial information) and *dynamic selection* (based on additional information as well). Understanding this distinction is crucial for assessing whether PT violations arise from mechanisms we are interested in, or from factors we do not intend to address.

Selection into treatment at each decision point follows different mechanisms:

- when choosing D_0, selection can only depend on the initial information set U_0. This follows directly from the choice rule for D_0, which we previously derived. As a result, selection into treatment at $t = 0$ is necessarily static—it does not involve any dynamic mechanisms;

- when choosing D_1, selection is influenced not only by the initial information U_0 but also by any additional information obtained by observing the first-period outcome $Y_0(d_0)$. Consequently, selection into treatment at $t = 1$ can be static (if it depends only on U_0) or dynamic (if it incorporates new information from period 0).

This distinction between static and dynamic selection is relevant because we wish not to account for static selection when evaluating PT. The reason is that, absent a pre-treatment period, static selection is unavoidable. To do this, we introduce an additional assumption.

Assumption (11). The untreated trend is mean-independent of the initial information set U_0, such that:

$$\mathbb{E}[Y_1(0) - Y_0(0) \mid U_0] = \tau. \tag{11}$$

This assumption safeguards PT from static selection: if treatment decisions (D_0, D_1) are fully determined by U_0, PT remains valid under this assumption. However, PT may still be violated if treatment at $t = 1$ is determined based on new information revealed by observed outcomes.

With these foundations in place, we now turn to the central question about the relationship between learning and identification. We analyze three distinct scenarios, each illustrating a different way in which information acquisition influences treatment selection.

3.1 Sufficient initial information.

We begin our analysis by considering a setting in which the initial information U_0 is *sufficient to determine mean outcomes*. In other words, agents have complete knowledge of the expected potential outcomes, as there is no scope for learning over time. The key question is whether PT holds under these conditions. While we have already suggested that it might, we now formalize this intuition.

Example 2. Suppose DUM and assumption (11) hold. If the initial information is sufficient for mean outcomes in the following sense:

$$\mathbb{E}[Y_1(D_1) \mid U_0, Y_0(d_0)] = \mathbb{E}[Y_1(D_1) \mid U_0], \tag{12}$$

the PT condition is satisfied. Assumption (12) implies that observing $Y_0(d_0)$ does not provide any additional information about $Y_1(D_1)$ beyond what is already contained in U_0. Since no learning occurs, treatment decisions remain entirely determined by initial information.

Proof. In the DUM model we have the following choice rule for D_1:

$$\begin{aligned} D_1 &= 1\big\{E[V_1(d_0,d_1) - V_1(d_0,0)|U_0,U_1(d_0)] \geq 0\big\} \\ &= 1\big\{E[V_1(d_0,d_1) - V_1(d_0,0)|U_0] \geq 0\big\}, \text{per assumption (12).} \end{aligned}$$

That implies D_1 is conditional on U_0 only, like D_0. Thus, by the law of iterated expectations (LIE) and assumption (11) – which states that $\mathbb{E}[Y_1(0) - Y_0(0) \mid U_0]$ is a constant τ –, the conditional untreated trend is:

$$\begin{aligned} E[Y_1(0) - Y_0(0)|D_0,D_1] &= E\big[E[Y_1(0) - Y_0(0)|U_0]\big|D_0,D_1\big] \\ &= E\big[\tau|D_0,D_1\big] = \tau, \text{ for all } (d_0,d_1). \end{aligned}$$

As a result, PT is valid.

Interpretation. Intuitively, this result follows because the treatment decision D_1 does not introduce additional selection mechanisms beyond those already present in U_0. Since we have assumed that static selection does not invalidate PT, and no new selection occurs at $t = 1$, we can conclude that PT is valid under these conditions.

To understand the implications of this result in a real-world setting, consider our job training example. When initial information is sufficient for predicting future wages, workers cannot improve the accuracy of their expectations over time. Consequently, they make treatment decisions solely based on their pre-existing beliefs at $t = 0$. This leads to static self-selection into treatment, meaning individuals choose whether to enroll based on their perceived best option at the start of the study.

Crucially, thanks to Assumption (11), static selection does not violate PT. In this scenario, workers do not update their expectations after observing their initial wage outcomes, meaning that $U_1(d_0)$ does not provide any additional information beyond U_0. Without new information, there is no dynamic treatment selection, and PT holds.

While this setting establishes a useful benchmark, it is somewhat idealized. In practice, individuals may have static expectations about their untreated wages – perhaps assuming they remain roughly constant over time –, but uncertain expectations about the returns to treatment. For instance, workers may have a good sense of how their wages will evolve in the absence of training but lack precise knowledge about how the training program will affect their future earnings.

This situation resembles a multi-armed bandit problem, where workers face a choice between a *safe arm* (remaining untreated, with predictable wages) and a *risky arm* (enrolling in training, with uncertain returns). When learning is possible but limited to treated outcomes, the decision-making process becomes more complex. In the next section, we examine how learning about treatment effects – rather than learning about untreated potential outcomes – impacts PT.

3.2 Learning on the Treatment arm.

We now extend the previous setting by introducing learning that is *restricted to treated outcomes*. That is, individuals can update their beliefs about the effects of treatment, but they do not acquire additional information about untreated outcomes. This scenario generalizes the previous case: whereas in the first case, the initial information U_0 was sufficient for both treated and untreated outcomes, we now assume it is only sufficient for untreated outcomes.

Example 3. Suppose DUM and assumption (11) hold. If:

- there is no learning across treatment arms ($d_0 \neq d_1$) nor on the untreated arm ($d_0 = d_1$):

$$E[Y_1(d_1)|U_0, Y_0(d_0)] = E[Y_1(d_1)|U_0] \text{ where } (d_0 + d_1) \in \{0, 1\}, \tag{13}$$

- $Y_0(1)$ is not informative about $Y_0(0)$ given U_0:

$$E[Y_0(0)|Y_0(1), U_0] = E[Y_0(0)|U_0] \text{ for any } d_0, d_1 \in \{0, 1\}, \tag{14}$$

the PT condition is satisfied.

Proof. We first analyze the group of individuals who are not treated at $t = 0$, irrespective of whether they later receive treatment at $t = 1$. Since D_0 is determined solely by U_0, we can condition on U_0 alone. However, this is not necessarily true for D_1, since treatment decisions at $t = 1$ may depend on new information.

A key observation, however, is that the second-period treatment choice of individuals who remain untreated at $t = 0$, $D_1(0)$, must also be determined solely by U_0. The reasoning follows from assumption (14): since no new information is acquired when not treated at $t = 0$, the information set is the same as U_0. Applying the LIE, we condition on $D_0 = 0$ and $D_1(0)$, leading to:

$$\begin{aligned} E[Y_1(0) - Y_0(0)|D_0 = 0, D_1 = d_1] \\ = E\big[E[Y_1(0) - Y_0(0)|U_0]\big|D_0 = 0, D_1(0) = d_1\big] \\ = E\big[\tau|D_0 = 0, D_1(0) = d_1\big] \\ = \tau, \text{ for all } d_1 \in \{0, 1\}. \end{aligned}$$

By assumption (11), the untreated trend is constant given U_0. Since we have successfully isolated it, we conclude that the untreated trend among those untreated at $t = 0$ is constant.

We now turn to individuals who were treated at $t = 0$ and consider their second-period treatment choice $D_1(1)$. Unlike in the previous case, here $D_1(1)$ depends not only on U_0 but also on the realized outcome $Y_0(1)$. This is because individuals in this group have received treatment and, according to assumption (13), can update their beliefs about treated potential outcomes. By the LIE:

$$\begin{aligned} &E[Y_1(0) - Y_0(0)|D_0 = 1, D_1(1) = d_1] \\ = {} &E\big[E[Y_1(0) - Y_0(0)|U_0, Y_0(1)]\big|D_0 = 1, D_1(1) = d_1\big]. \end{aligned}$$

At first glance, this may seem problematic for PT, as it introduces a dependency on an observed outcome. However, assumption (14) states that no learning about untreated outcomes occurs. This has a crucial implication: although $D_1(1)$ depends on $Y_0(1)$, this dependency does not affect expectations about untreated potential outcomes.

We can thus ignore the conditioning on $Y_0(1)$ and simplify by assumption (11):

$$\begin{aligned} &E\big[E[Y_1(0) - Y_0(0)|U_0, Y_0(1)]\big|D_0 = 1, D_1(1) = d_1\big] \\ = {} &E\big[E[Y_1(0) - Y_0(0)|U_0]\big|D_0 = 1, D_1(1) = d_1\big] = \tau, \text{ for all } d_1 \in \{0, 1\}. \end{aligned}$$

Thus, the untreated trend among those treated at $t = 0$ is also constant and takes the same value τ dictated by assumption (11). This confirms that learning about treated outcomes does not affect the constancy of untreated trends.

With these results in place, we now verify that PT is satisfied by showing that the expected untreated trend is independent of treatment assignment in both periods.

1. We have shown that the conditional untreated trend is constant in d_1 for any given d_0, corresponding to statement (i) of Observation 1, ie. there is no selection in choosing D_1.

2. We have shown that the untreated trend conditional on d_0 is constant in d_0, corresponding to statement (ii), ie. there is no selection in choosing D_0.

3. Thanks to assumption (11), we can see that the untreated trend always takes the same value τ, corresponding to statement (iii), ie. there is no trend shift between the two periods.

These results establish that Observation 1 is satisfied, thus PT holds in this setting.

Interpretation. This case highlights an important implication: dynamic treatment selection does not necessarily violate PT. If agents only learn about treatment effects and not about untreated potential outcomes, PT is preserved. However, if learning extends to untreated outcomes, the situation may change.

In the next section, we examine a case where learning occurs across both treatment and control groups, and we explore the consequences for identification.

3.3 Learning on the Control arm.

In the last scenario, we examine the case where learning occurs *on untreated outcomes* rather than treated outcomes. We assume that U_0 is sufficient for accurately predicting treated outcomes, but individuals don't know how their untreated outcomes will evolve over time. This implies that individuals who are not treated at $t = 0$ can acquire new information about untreated outcomes, which may influence their subsequent decisions.

Example 4. Suppose DUM and assumption (11) hold. Let Ω_{vl} be the set of *valuable learners*, ie. agents whose optimal D_1 is decisively influenced by Y_0. If $Y_0(0)$ is informative about $Y_1(0)$ but no other form of learning is possible, then PT is satisfied if and only if:

- valuable learning is impossible (ie. $P(\Omega_{vl}) = 0$), or
- valuable learning occurs only where untreated outcomes are constant almost surely (ie. $P(Y_0(0) = Y_1(0)|\Omega_{vl}) = 1$) and $\tau = 0$.

Under these conditions, we generally expect PT not to be satisfied. However, it remains valid under specific circumstances, which we examine below. Before delving into these exceptions, we first explore why identification is undermined here, without formally proving it.

Justification. To understand the breakdown of PT, we revisit the proof structure from the second case. There, we established that the untreated trend was constant among those untreated at $t = 0$, ensuring that PT remained valid. The same reasoning applied to individuals treated at $t = 0$ since there was no learning about untreated outcomes, meaning that treatment status at $t = 1$ did not introduce new selection biases. However, in this case, the proof fails precisely when conditioning on individuals who remain untreated at $t = 0$. Unlike in the previous setting, those subjects do update their expectations in this case. Since those who are treated at $t = 0$ do not gain new information, their expectations remain unchanged. This asymmetric information update creates systematic differences between individuals who receive early treatment and those who do not.

Although learning on untreated outcomes typically invalidates PT, there exist two conditions under which it remains intact.

The first condition is that there is no *valuable* learning taking place. Learning is defined as valuable if it is decisive for treatment decisions at $t = 1$. If agents do not update their expectations in a way that alters their second-period treatment choice, then learning has no impact on selection into treatment, and PT is preserved. In this case, observing the first outcome does not induce differential selection across groups, ensuring that the untreated trend remains constant.

The second condition arises when valuable learning does take place, but only when untreated potential outcomes are stable over time. If untreated outcomes do not change across periods, then the untreated trend is necessarily zero, trivially satisfying the PT condition.

Interpretation. This analysis demonstrates that learning on untreated outcomes introduces a fundamental asymmetry between individuals who remain untreated and those who receive early treatment. Unlike in previous cases, individuals who delay treatment gain new information about their potential outcomes under non-treatment, which can lead to dynamic selection effects that invalidate PT. However, PT is preserved in cases where learning does not alter treatment decisions or when untreated outcomes remain constant over time.

These findings underscore the bottom line of this exercise: learning does not necessarily undermine PT – the type of learning matters. When agents update beliefs only about treated outcomes, PT remains valid. When learning extends to untreated outcomes, PT generally fails. This distinction is crucial for empirical applications, where researchers must assess not only whether learning occurs, but also how it influences treatment selection dynamics.

In the next section, we introduce a simulation to demonstrate the findings of this scenario and to display the main ideas of this section.

4. Simulation

We mimic the setting we introduced in the Setup section. We have a sample of N workers for whom potential (monthly) wages at time $t \in \{0, 1\}$ are given by the following equation:

$$Y_{it}(d_{it}) = \alpha + \lambda \cdot t + \tau \cdot d_{it} + e_{it}.$$

α is a (common) baseline that may reflect, for instance, the statutory minimum wage, λ is the (common) linear time trend that may reflect inflation-driven wage growth, and τ is the ATT, ie. the average effect of the job training program on the wages of workers who participated in it. We assume that the effect is constant over time, homogeneous for all workers (reflecting perfect compliance and no intrinsic effect heterogeneity), and independent from other variables. e_{it} is an unobservable variable that shall reflect the effort of worker i in period t.

Treatment d_{it} is determined as in the DUM model illustration presented in subsection 2.2, ie. treatment decisions are made sequentially by individual workers on the basis of their current information U_{it}. U_{it} comprises:

- the (common) discount rate β, with $\beta = 0.5$ here;
- the (common) program cost structure $K_t(\cdot)$;
- the baseline α, with $\alpha = 1200$ here;
- the time trend λ, with $\lambda = 100$ here;
- individual effort e_{it}, known to worker i though not observable by the investigator;
- individual expectations or predictions $\hat{\tau}_{it}$ on the treatment effect.

In our setting, enrolling first in $t = 0$ is expensive enough, with $K_0(1) = 7000$, to force all workers to remain untreated in the first period, while enrollment costs are equal to $K_1(0, 1) = K_1(1, 1) = 700$ in $t = 1$, regardless of whether it is the second or first enrollment for the individual worker. Not enrolling is always free, ie. $K_0(0) = K_1(0, 0) = K_1(1, 0) = 0$. We want to force a sharp design to keep the example as simple as possible: that implies no asymmetric selection in the first period and, thus, no need to satisfy assumption (11) since Parallel Trends can be violated solely by dynamic mechanisms. It follows that the conventional DiD estimator can be used to estimate the ATT in an unbiased fashion, absent dynamic selection.

Individual effort in the first period is given by $e_{i0} \overset{\text{iid}}{\sim} \mathcal{N}(\mu_\epsilon, \sigma_\epsilon)$, with mean $\mu_\epsilon = 1200$ and standard deviation $\sigma_\epsilon = 300$ here. After observing $Y_{i0}(d_{i0}) = Y_{i0}$, each worker determines a new effort $e_{i1} = (1 + \rho) \cdot e_{i0}$, where ρ expresses the learning rate. The rationale is similar to the one adopted by Fudenberg & Levine, 2022: by observing their realized outcomes, workers learn on their own effort, or – in other words – they learn about the effect of their efforts. This process is modelled in a simplistic and relatively unstructured fashion, since it implies a positive feedback loop where the relative effort increase is constant for all workers. Nonetheless, it is parsimonious and serves our purpose well.

As anticipated, treatment status in $t = 0$ is determined by maximizing the expected or predicted discounted utility:

$$\hat{U}_{i0}(d_{i0}) = \hat{Y}_{i0}(d_{i0}) - K_0(d_{i0}) + \beta \cdot (\hat{Y}_{i1}(d_{i1}) - K_1(d_{i0}, d_{i1}))$$

over $d_{it} \in \{0, 1\}$, where $\hat{Y}_{it}(d_{it}) = \alpha + \lambda \cdot t + \hat{\tau}_{it} \cdot d_{it} + e_{it}$. Treatment status in $t = 1$ is determined analogously, though without the forward-looking component, by maximizing:

$$\hat{U}_{i1}(d_{i0}, d_{i1}) = \hat{Y}_{i1}(d_{i1}) - K_1(d_{i0}, d_{i1})$$

over $d_{i1} \in \{0, 1\}$.

Let us start with $\rho = 0.5$. We can see the observed DiD plot in Figure 2. Unexpectedly, the switchers into treatment enjoy a stronger wage growth.

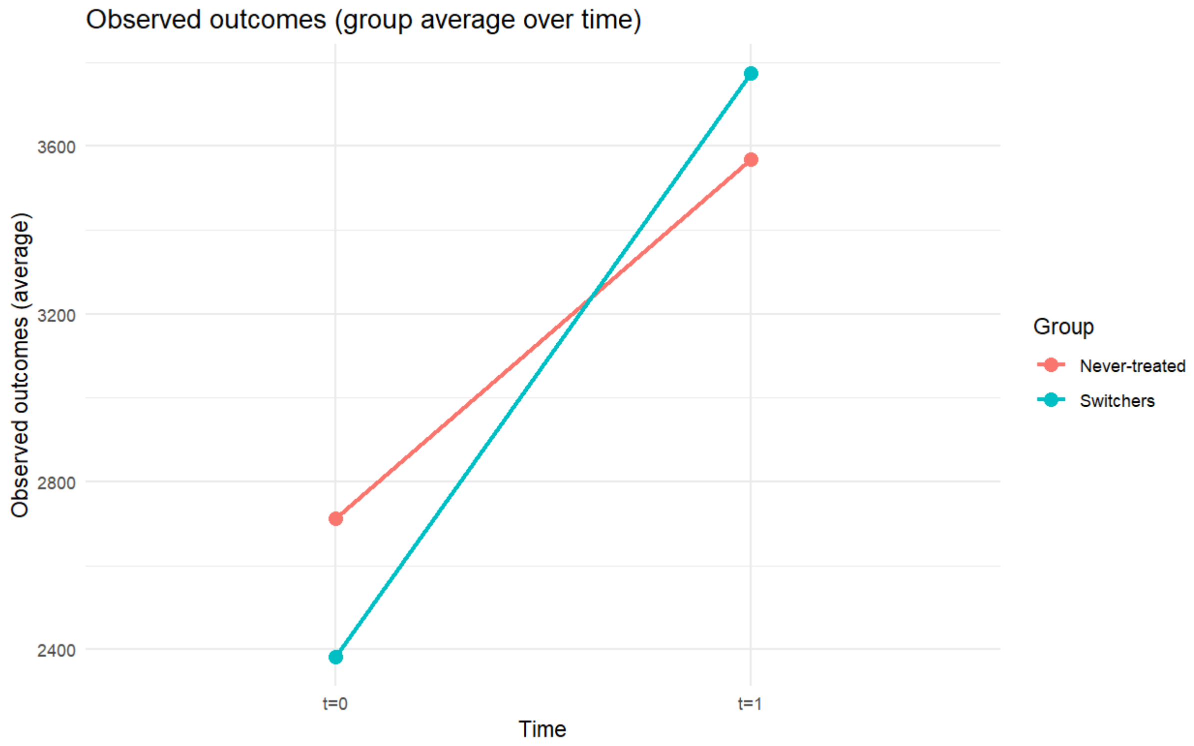

Figure 2: Observed DiD plot with increasing effort.

In Table 1 we can see how the confidence intervals (estimated via bootstrap) for $\hat{\tau}_{DiD}$ change over different learning rates. The estimates are consistently biased, curiously in either direction. But the bias seemingly gets smaller for decreasing values of ρ, at least for $\rho < 0.1$. As the learning effect decreases, the bias tends to decrease as we would expect.

Table 1: 95% Confidence Intervals for $\hat{\tau}_{DiD}$

ρ	Lower bound	Upper bound
0	700.0	700.0
0.001	700.5	700.6
0.005	702.7	703.1
0.01	705.3	706.3
0.05	718.1	727.7
0.1	717.1	743.5
0.5	523.8	546.8
1.0	669.5	688.4

For $\rho = 0.5$, the conventional DiD estimator returns $\hat{\tau}_{DiD} = 534.4$, an evidently biased estimate of the ATT. In fact, in Figure 3 we can see that PT is not satisfied in this case, since there is a considerable difference in the untreated trends between the two groups.

Lastly, in Figure 4 we can see the untreated trends for $\rho = 0.01$, a neglectable learning rate. The PT assumption can be safely assumed if learning effect is reasonably small.

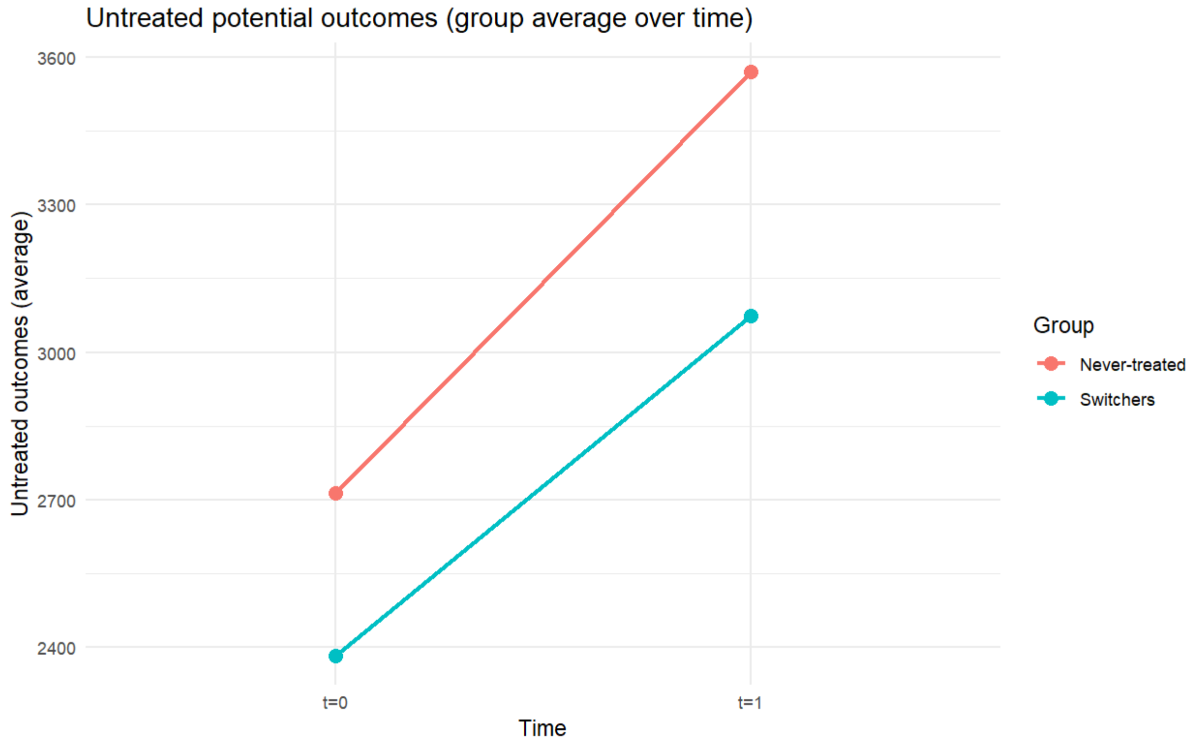

Figure 3: Counterfactual DiD plot with increasing effort.

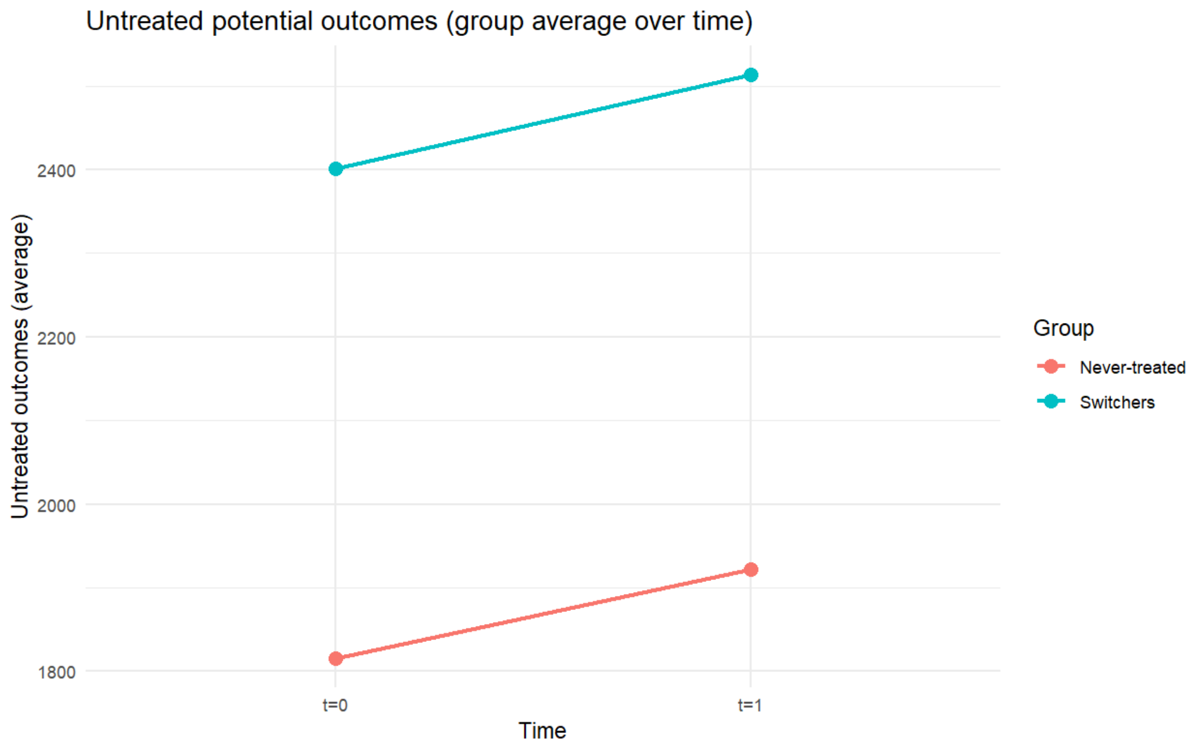

Figure 4: Counterfactual DiD plot with neglectable increase in effort.

Note that learning on effort does not alter exclusively untreated outcomes, in fact it alters all potential outcomes in the second period. But it can be shown that PT would not be violated if learning was limited to updating the predictions $\hat{\tau}_{it}$.

5. Conclusions

Since Parallel Trends constrain agent behavior, the assumption is not necessarily compatible with dynamic choice mechanisms. Our examples specifically deal with learning about potential outcomes.

Example 2. PT may hold if observed outcomes do not provide additional information about the future, ie., if agents do not learn anything new about potential outcomes. *No dynamic selection in this case.*

Example 3. PT may hold if agents can learn only from and about treated outcomes, ie., if untreated outcomes are accurately expected from the start. *Learning on treated outcomes does not affect untreated trends.*

Example 4. PT may not hold if valuable information can be learned from and about untreated outcomes, unless said outcomes are stable over time. *Learning on untreated outcomes may lead to differences in conditional untreated trends.*

Acknowledgments

I warmly thank Prof. Daniel Wilhelm for his supervision in the *Machine Learning in Econometrics* Seminar (2024/25) and his helpful feedback. I also wish to thank my seminar colleagues for their inspiring work and the motivation they gave me. I acknowledge the use of ChatGPT for background research, conceptual clarification, literature review, coding suggestions, refinement of arguments, and editorial improvements. All decisions, findings, interpretations, conclusions, and any remaining errors remain solely my responsibility.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, 72(1), 1–19. https://doi.org/10.1111/0034-6527.00321

Abbring, J. H. (2010). Identification of dynamic discrete choice models. *Annual Review of Economics*, 2(Volume 2, 2010), 367–394. https://doi.org/https://doi.org/10.1146/annurev.economics.102308.124349

Arkhangelsky, D., & Imbens, G. (2024). Causal models for longitudinal and panel data: A survey. *The Econometrics Journal*, 27(3), C1–C61. https://doi.org/10.1093/ectj/utae014

Arkhangelsky, D., & Imbens, G. W. (2022). Doubly robust identification for causal panel data models. *The Econometrics Journal*, 25(3), 649–674. https://doi.org/10.1093/ectj/utac019

Ashenfelter, O., & Card, D. (1985). Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *The Review of Economics and Statistics*, 67(4), 648–660.

Ban, K., & Kédagni, D. (2023). Robust difference-in-differences models. https://arxiv.org/abs/2211.06710

de Chaisemartin, C., & D'HaultfŒuille, X. (2017). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2), 999–1028. https://doi.org/10.1093/restud/rdx049

de Chaisemartin, C., D'Haultfœuille, D. C. X., & Knau, F. (2024). Two-way fixed effects and differences-in-differences estimators in heterogeneous adoption designs. https://arxiv.org/abs/2405.04465

Dobkin, C., Finkelstein, A., Kluender, R., & Notowidigdo, M. J. (2018). The economic consequences of hospital admissions. *American Economic Review*, 108(2), 308–52. https://doi.org/10.1257/aer.20161038

Freyaldenhoven, S., Hansen, C., & Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9), 3307–38. https://doi.org/10.1257/aer.20180609

Fudenberg, D., & Levine, D. K. (2022). Learning in games and the interpretation of natural experiments. *American Economic Journal: Microeconomics*, 14(3), 353–77. https://doi.org/10.1257/mic.20200106

Ghanem, D., Sant'Anna, P. H. C., & Wüthrich, K. (2024). Selection and parallel trends. https://arxiv.org/abs/2203.09001

Heckman, J. J., & Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects [The interface between econometrics and economic theory]. *Journal of Econometrics*, 136(2), 341–396. https://doi.org/https://doi.org/10.1016/j.jeconom.2005.11.002

Lechner, M., et al. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics*, 4(3), 165–224.

Manski, C. F., & Pepper, J. V. (2018). How do right-to-carry laws affect crime rates? coping with ambiguity using bounded-variation assumptions. *Review of Economics and Statistics*, 100(2), 232–244.

Marx, P., Tamer, E., & Tang, X. (2024). Parallel trends and dynamic choices. *Journal of Political Economy Microeconomics*, 2(1), 129–171.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218–2244.

Taber, C. R. (2001). The rising college premium in the eighties: Return to college or return to unobserved ability? *The Review of Economic Studies*, 68(3), 665–691. https://doi.org/10.1111/1467-937X.00185

Verdier, V. (2020). Average treatment effects for stayers with correlated random coefficient models of panel data. *Journal of Applied Econometrics*, 35(7), 917–939. https://doi.org/https://doi.org/10.1002/jae.2789

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70(1), 331–341. https://doi.org/https://doi.org/10.1111/1468-0262.00277

Appendix

The R script that was used for the simulation can be found on GitHub. The code is fully reproducible with seed 153. It can also be easily adapted to a setting where learning happens only from and about treated outcomes, for instance if workers update their predictions about the treatment effect while their effort is constant, or in more complex settings.