Research Article

Markus Hersche and Elias Moor*

Identification and Estimation of Intensive Margin Effects by Difference-in-Difference Methods

<https://doi.org/10.1515/jci-2019-0035>

Received Jun 03, 2019; accepted Oct 01, 2020

Abstract: This paper discusses identification and estimation of causal intensive margin effects. The causal intensive margin effect is defined as the treatment effect on the outcome of individuals with a positive outcome irrespective of whether they are treated or not, and is of interest for outcomes with corner solutions. The main issue is to deal with a potential selection problem that arises when conditioning on positive outcomes. We propose using difference-in-difference methods - conditional on positive outcomes - to estimate causal intensive margin effects. We derive sufficient conditions under which the difference-in-difference estimator identifies the causal intensive margin effect. We apply the methodology to estimate the causal intensive margin effect of reaching the full retirement age on working hours.

Keywords: Intensive margin effect, difference-in-difference, corner solutions, policy evaluation

2020 Mathematics Subject Classification: 62D20, 62P20

1 Introduction

A decomposition of a binary treatment into extensive and intensive margin effects is of special interest when studying outcomes with a corner solution at zero.¹ Outcomes with corner solutions include working hours, health expenditures, and trade volumes. The average effect of a treatment on an outcome with a corner solution at zero can be decomposed into 1) the average change in the outcome of those with a positive outcome irrespective of treatment, plus 2) the average outcome of those with a positive outcome in case of treatment and a zero outcome in case of no treatment, minus 3) the average outcome of those with a zero outcome in case of treatment and a positive outcome in case of no treatment [1–3]. Part 1) represents the weighted causal intensive margin effect. The sum of 2) and 3) captures the weighted causal extensive margin effect.²

Article note: We are thankful to Antoine Bommier, Rainer Winkelmann, Michael Lechner, Alexis Direr, and Tobias Wekhof, as well as conference participants at the 2018 French Econometrics Conference for their helpful comments. An earlier version of this paper was part of the doctoral thesis of Markus Hersche (“Theoretical and Empirical Essays on Labor Supply of the Elderly”, *Diss. ETH No. 25377*, <https://doi.org/10.3929/ethz-b-000308517>) and Elias Moor (“Essays on Causal Inference in Economics: Methods and Applications”, *Diss. ETH No. 26810*, <https://doi.org/10.3929/ethz-b-000432494>), and was published as a working paper (“Identification of Causal Intensive Margin Effects by Difference-in-Difference Methods”, CER-ETH Economics Working Paper Series, 11/2018). We gratefully acknowledge financial support from the Swiss Re Foundation and the ETH Zurich Foundation.

***Corresponding Author: Elias Moor:** Chair of Integrative Risk Management and Economics, ETH Zurich; Email: emoor@ethz.ch

Markus Hersche: Chair of Integrative Risk Management and Economics, ETH Zurich; Email: mhersche@ethz.ch

1 Corner solutions at alternative thresholds are possible as well. For simplicity and illustration, we consider the case where the threshold is at zero.

2 The weights are given by the relative size of the group in the population.

Take as an example the effect of the introduction of a partial retirement policy on labor supply. The aim of such a policy is to increase labor market participation of older people and to preserve human capital. Suppose that in the status quo, individuals must withdraw the full pension at a given age, but are allowed to continue working.³ Under the partial retirement policy, individuals have the choice between a partial and a full pension, and are allowed to continue working. Claiming a partial pension can be attractive because it increases subsequent pension benefits as a result of an extended contribution period as well as reduced and delayed pension claims. The total effect of such a policy on labor supply might be zero, suggesting that the policy has been ineffective. The zero result, however, could be explained by a positive extensive margin effect that was offset by a negative intensive margin effect. Older workers who would have retired in the absence of a partial retirement policy, may now decide to stay in the labor market. Likewise, individuals who would have worked full-time in the absence of a partial retirement policy, may now decide to work part-time. In such cases the total effect masks interesting subeffects at the extensive and intensive margin.

Even if treatment is randomly assigned, estimating intensive margin effects is challenging. A mean comparison of treatment and control groups with positive outcomes does not identify the causal intensive margin effect without additional assumptions [4]. In the labor supply example, the sample of individuals with positive working hours consists of two groups: 1) the group of individuals with positive working hours irrespective of whether they are treated or not (always-participants), i.e. irrespective of whether they have the possibility to withdraw a partial pension; and 2) the group of individuals with positive working hours only because they are treated (switchers), i.e. only because they have the possibility to withdraw a partial pension, who would not work if they could only withdraw the full pension.⁴ For the causal intensive margin effect, we are only interested in the group of always-participants. Group membership is however not observed in the data, because we observe either the outcome in case of treatment or the outcome in case of no treatment. The unobserved characteristics of always-participants and switchers are likely to be different. Always-participants might be more motivated than switchers. Therefore, average working hours of always-participants are likely higher than average working hours of switchers. As a result, a difference in the means of treated and untreated – conditional on positive working hours – could be the result of differences in these unobserved characteristics, and not because of a causal effect of treatment.

This constitutes a selection problem. In a general setting without random treatment, we are thus faced with two selection problems. The first selection problem is the standard selection problem in observational studies. In the presence of confounding variables, a mean comparison of treated and untreated individuals does not identify the causal effect. The second selection problem arises because we condition on positive outcomes. Difference-in-difference methods were developed to deal with the first selection problem. Using data from pre- and post-treatment periods, difference-in-difference allows for some selection on unobservables. This comes at the cost of making an assumption about outcome trends over time. It seems reasonable to extend the difference-in-difference methodology to include the second selection problem as well.

In this paper, we introduce difference-in-difference methods to estimate the causal intensive margin effect. Compared to standard difference-in-difference estimators, we condition the sample on individuals with positive outcomes.⁵ We derive sufficient conditions under which the causal intensive margin effect is identified. In contrast to standard difference-in-difference methods, two monotonicity assumptions are additionally required to identify the causal intensive margin effect. We apply the difference-in-difference methodology to estimate the causal intensive margin effect of reaching the full retirement age on working hours. Knowledge of the causal intensive margin effect is important to obtain a better understanding of retirement choices, for

³ That means we consider a setting without mandatory retirement.

⁴ Here we neglect a possible third group, the group of individuals with positive working hours only because they are not treated, i.e. only because they *do not* have the possibility to withdraw a partial pension, and who would not work if they had the choice between a partial and a full pension. In this example this case seems rather unlikely.

⁵ We refer to the term *standard difference-in-difference* to denote difference-in-difference methods that do not condition on positive outcomes.

example to improve the design of future pension reforms. Moreover, we discuss how the identifying assumptions can be motivated in practice.

The main contribution of this paper is to extend the literature on identification and estimation of intensive margin effects by borrowing well established difference-in-difference methods from the policy evaluation literature. The intensive margin effect is of interest in cases where the total effect masks relevant subeffects, e.g. when the extensive and intensive margin effect have different signs. Compared to models for outcomes with corner solutions or selection models, the difference-in-difference estimator on positive outcomes is based on different assumptions; most importantly on a common trend assumption.

This paper is related to the literature on models for outcomes with corner solutions, e.g. Tobit [5, 6] or two-part models [7, 8], and selection models [9]. Moreover, the paper is connected to the literature employing principal stratification following [10] to study causal extensive and intensive margin treatment effects for variables with nonnegative outcomes [1–3]. This literature decomposes the average treatment effect into a population-weighted sum of treatment effects on always-participants and switchers. Studying outcomes with a corner solution at zero, [2] derives nonparametric bounds for the treatment effects on always-participants and switchers. He further discusses point identification of causal intensive and extensive margin effects in censored regression, selection, and two-part models. [1, 3] analyzes total, extensive, and intensive margin effects in general sample selection models, with the corner solution outcome as a special case. [1] analyzes nonparametric methods to estimate extensive and intensive margin effects, whereas [3] discusses point identification of intensive and extensive margin effects in semiparametric linear models. The notion of principal stratification is also used in instrumental variable approaches [11], and in mediation analysis [12]. In instrumental variable approaches, the stratification is based on the treatment variable (always-takers, compliers, defiers, never-takers), whereas in mediation analysis, the stratification is based on the mediator. In our context, the stratification is based on the outcome variable. More generally, the paper often draws upon [13], who provides a survey on difference-in-difference methods from a potential outcomes perspective.

The remainder of the paper is organized as follows. Section 2 introduces the notation and describes the conventional as well as the causal decomposition of a treatment effect. Identification of the causal intensive margin effect is described in Section 3. Section 4 discusses estimation and inference. An empirical application is presented in Section 5. The last section concludes.

2 Notation and Decomposition of a Treatment Effect

2.1 Notation

We consider the standard potential outcome framework with a non-negative outcome Y and an indicator for the treatment group D [14], extended to two periods [13]. We observe individuals in the pre-treatment period $t - 1$, and in the post-treatment period t ; that is we observe $Y_{i,t-1}$ and $Y_{i,t}$. In each period, each individual i has two potential outcomes. The potential outcomes in case of treatment ($D_i = 1$) are denoted by $Y_{i,t}^1$ and $Y_{i,t-1}^1$, and in case of no treatment ($D_i = 0$) by $Y_{i,t}^0$ and $Y_{i,t-1}^0$.⁶ In each period, we only observe one of the two potential outcomes.

Moreover, each individual is characterized by a vector of observed covariates X_i , assumed to be constant over time. The starting point of the decompositions described in Sections 2.2 and 2.3 is the average treatment effect on the treated (ATT), defined as

$$ATT_t = E(Y_{i,t}^1 - Y_{i,t}^0 | D_i = 1). \quad (1)$$

The ATT measures the expected treatment effect for a treated observation. Hence, this specification allows for arbitrary treatment effect heterogeneity.

⁶ Note that the treatment indicator is not indexed with a time index, i.e. $D_i = 1$ for individuals treated between $t - 1$ and t , and $D_i = 0$ for individuals not treated between $t - 1$ and t .

2.2 Conventional decomposition

As described in Section 1, the estimation of causal intensive margin effects entails two selection problems. The first selection problem arises from confounding variables, the second selection problem arises from conditioning on observations with positive outcomes. To illustrate the second selection problem, we consider in this subsection the case of random treatment assignment, and thus eliminate the first selection problem. This illustration closely follows [2]. Random treatment assignment implies that treatment is independent of the potential outcomes, i.e. $(Y_{i,t}^1, Y_{i,t}^0) \perp\!\!\!\perp D_i$. Hence, the ATT at time t is identified by the difference in mean outcomes of treated and untreated:⁷

$$ATT_t = E(Y_{i,t}^1 | D_i = 1) - E(Y_{i,t}^0 | D_i = 1) \quad (2)$$

$$= E(Y_{i,t} | D_i = 1) - E(Y_{i,t} | D_i = 0) \quad (3)$$

A non-negative outcome (with a point mass at zero) is often decomposed into an extensive and an intensive part as $E(Y_{i,t}) = E(Y_{i,t} | Y_{i,t} > 0)P(Y_{i,t} > 0)$. Similar to [2], the difference in mean outcomes can then be rewritten as

$$ATT_t = E(Y_{i,t} | D_i = 1) - E(Y_{i,t} | D_i = 0) \quad (4)$$

$$= E(Y_{i,t} | Y_{i,t} > 0, D_i = 1)P(Y_{i,t} > 0 | D_i = 1) \quad (5)$$

$$- E(Y_{i,t} | Y_{i,t} > 0, D_i = 0)P(Y_{i,t} > 0 | D_i = 0)$$

$$= \underbrace{[P(Y_{i,t} > 0 | D_i = 1) - P(Y_{i,t} > 0 | D_i = 0)] E(Y_{i,t} | Y_{i,t} > 0, D_i = 1)}_{\text{extensive margin effect}} \quad (6)$$

$$+ \underbrace{[E(Y_{i,t} | Y_{i,t} > 0, D_i = 1) - E(Y_{i,t} | Y_{i,t} > 0, D_i = 0)] P(Y_{i,t} > 0 | D_i = 0)}_{\text{intensive margin effect}}. \quad (7)$$

The terms in (6) represent the extensive margin effect, the terms in (7) the intensive margin effect. Under random treatment, the terms in (6) and (7) can be rewritten as

$$ATT_t = [P(Y_{i,t}^1 > 0) - P(Y_{i,t}^0 > 0)] E(Y_{i,t}^1 | Y_{i,t}^1 > 0) \quad (8)$$

$$+ [E(Y_{i,t}^1 | Y_{i,t}^1 > 0) - E(Y_{i,t}^0 | Y_{i,t}^0 > 0)] P(Y_{i,t}^0 > 0). \quad (9)$$

The difference in (8) is a causal comparison and captures the causal effect of treatment on the probability of having a positive outcome. However, the difference in (9) does generally not have a causal interpretation, because we compare two possibly different subgroups of the population. The subgroup with a positive outcome in case of treatment ($Y_{i,t}^1 > 0$) and the subgroup with a positive outcome in case of no treatment ($Y_{i,t}^0 > 0$). Hence, conditioning on positive outcomes induces a selection problem. As a result, the difference in mean outcomes of treated and untreated – conditional on positive outcomes – does not identify the causal intensive margin effect (without additional assumptions, see Appendix A). In the next section, we use a decomposition in which both the extensive and the intensive part have a causal interpretation.

2.3 Causal Decomposition

Following [1] and [2], we define four exhaustive and mutually exclusive subgroups based on the joint distribution of potential outcomes in period t :

⁷ In the case of random treatment assignment, the average treatment effect on the treated is equal to the average treatment effect, defined as $ATE_t = E(Y_{i,t}^1 - Y_{i,t}^0)$.

Table 1: Subgroups Based on the Joint Distribution of Potential Outcomes in Period t

	$Y_{i,t}^0 = 0$	$Y_{i,t}^0 > 0$
$Y_{i,t}^1 = 0$	never-participants	switchers 2
$Y_{i,t}^1 > 0$	switchers 1	always-participants

Based on this definition, we decompose the average treatment effect on the treated (ATT) at time t as follows:

$$ATT_t = E(Y_{i,t}^1 - Y_{i,t}^0 | D_i = 1) \quad (10)$$

$$= \underbrace{E(Y_{i,t}^1 | Y_{i,t}^1 > 0, Y_{i,t}^0 = 0, D_i = 1) P(Y_{i,t}^1 > 0, Y_{i,t}^0 = 0 | D_i = 1)}_{\text{weighted causal extensive margin effect (switchers 1)}} \quad (11)$$

$$- \underbrace{E(Y_{i,t}^0 | Y_{i,t}^1 = 0, Y_{i,t}^0 > 0, D_i = 1) P(Y_{i,t}^1 = 0, Y_{i,t}^0 > 0 | D_i = 1)}_{\text{weighted causal extensive margin effect (switchers 2)}} \quad (12)$$

$$+ \underbrace{E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1) P(Y_{i,t}^1 > 0, Y_{i,t}^0 > 0 | D_i = 1)}_{\text{weighted causal intensive margin effect (always-participants)}} \quad (13)$$

The terms in (11) and (12) represent the weighted causal extensive margin effect. The term in (11) describes the effect of treatment on the outcome of individuals with positive outcome in case of treatment and zero outcome in case of no treatment (switchers 1), weighted by the fraction of switchers 1. The term in (12) describes the effect of treatment on the outcome of individuals with zero outcome in case of treatment and positive outcome in case of no treatment (switchers 2), weighted by the fraction of switchers 2. The contribution of individuals with zero outcome in the cases of treatment and no treatment (never-participants) is zero and therefore dropped.

The term in (13) represents the weighted causal intensive margin effect. It captures the effect of treatment on the outcome of individuals having a positive outcome irrespective of treatment status (always-participants), weighted by the fraction of always-participants.

In this decomposition, both the extensive margin effect and the intensive margin effect have a causal interpretation. In this paper we focus on the causal intensive margin average treatment effect on the treated.

3 Identification

We are interested in the intensive margin average treatment effect on the treated (IMATT),

$$IMATT_t = E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1) \quad (14)$$

$$= E \left[\underbrace{E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x)}_{\gamma_t(x)} \middle| Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1 \right]. \quad (15)$$

We will first derive sufficient conditions under which $\gamma_t(x)$, i.e. the conditional-on- X version of the intensive margin average treatment effect on the treated, is identified. In a second step, we state sufficient conditions under which the conditional-on- X version can be aggregated to $E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1)$.

3.1 Difference-in-Difference on Positive Outcomes

Difference-in-difference on positive outcomes is given by the difference of the time differences between treated and untreated observations

$$\begin{aligned} \gamma_t^{DiD}(x) = & E(Y_{i,t} - Y_{i,t-1} | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 1, X_i = x) \\ & - E(Y_{i,t} - Y_{i,t-1} | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 0, X_i = x). \end{aligned} \quad (16)$$

The following sufficient conditions identify the intensive margin average treatment effect on the treated.

Proposition 1 (Identification Difference-in-Difference on Positive Outcomes).

Sufficient conditions to identify the intensive margin average treatment effect on the treated using difference-in-difference on positive outcomes are

1. *stable unit treatment value assumption (SUTVA),*
2. *no pre-treatment effect,*
3. *common trend in positive outcomes,*
4. *no effect of treatment on covariates,*
5. *common support,*
6. *treatment monotonicity at the extensive margin, and*
7. *time monotonicity at the extensive margin.*

Assumptions 1-5 are also required in similar form in standard difference-in-difference. Assumptions 6 and 7 are specific to difference-in-difference on positive outcomes. These assumptions are additionally required to eliminate the selection problem arising from conditioning on individuals with positive outcomes. In the following we describe the assumptions in more detail.

Assumption 1 (SUTVA). *The stable unit treatment value assumption is given by*

$$\begin{aligned} Y_{i,t} &= (1 - D_i)Y_{i,t}^0 + D_i Y_{i,t}^1 \quad \forall i, \text{ and} \\ Y_{i,t-1} &= (1 - D_i)Y_{i,t-1}^0 + D_i Y_{i,t-1}^1 \quad \forall i, \end{aligned}$$

where $D_i \in \{0, 1\}$ denotes treatment status.

The SUTVA assumption ensures that we actually observe the potential outcomes in the treatment and control groups. The SUTVA assumption implies that the observed outcome of individual i only depends on the potential outcomes and the treatment status D_i , but not on the treatment status D_j of any other individual j . Thus, SUTVA rules out general equilibrium effects and spill-over effects.

Assumption 2 (No pre-treatment effect). *The no pre-treatment effect assumption is given by*

$$E(Y_{i,t-1}^1 - Y_{i,t-1}^0 | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 1, X_i = x) = 0 \text{ for all } x \text{ in the support of } X_i.$$

The no pre-treatment effect assumption requires that the treatment effect in the pre-treatment period is zero. Hence in expectation, individuals do not change their behavior in period $t - 1$ because they will be treated between period $t - 1$ and t .⁸

Assumption 3 (Common trend in positive outcomes). *The common trend in positive outcomes assumption is given by*

$$\begin{aligned} &E(Y_{i,t}^0 - Y_{i,t-1}^0 | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 1, X_i = x) \\ &= E(Y_{i,t}^0 - Y_{i,t-1}^0 | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 0, X_i = x) \text{ for all } x \text{ in the support of } X_i. \end{aligned}$$

The common trend in positive outcomes assumption represents the key assumption for identification. The common trend in positive outcomes assumption is closely related to the standard common trend assumption, except that we require the common trend to hold in the subsample of individuals with a positive outcome in period t and $t - 1$.⁹ The common trend in positive outcomes assumption requires that the treated and the

⁸ Using SUTVA, the no pre-treatment effect assumption can be rewritten to $E(Y_{i,t-1}^1 - Y_{i,t-1}^0 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x) = 0$, which is the version used in the proof of identification.

⁹ In standard difference-in-difference, the common trend assumption is given by $E(Y_{i,t}^0 - Y_{i,t-1}^0 | D_i = 1, X_i = x) = E(Y_{i,t}^0 - Y_{i,t-1}^0 | D_i = 0, X_i = x)$.

control group would experience the same time trend in case of no treatment.¹⁰ As [13] points out, the common trend assumption can be rewritten as a “constant bias” assumption. That is, the bias arising from unobserved confounders is assumed to be constant over time.

Assumption 4 (No effect of treatment on covariates). *The no effect of treatment on covariates assumption is given by*

$$X_i^1 = X_i^0 = X_i \quad \forall i.$$

The *no effect of treatment on covariates* assumption is required to ensure that conditioning on X does not condition away parts of the causal effect we are interested in, or introduce a collider bias.

Assumption 5 (Common support). *The common support assumption is given by*

$$P(D_i = 1 | Y_{i,t} > 0, Y_{i,t-1} > 0, X_i = x) < 1 \text{ for all } x \text{ in the support of } X_i.$$

The *common support* assumption requires that for all x in the support of X_i , there exist not only treated individuals in the subsample with positive outcomes in period t and $t - 1$.

Assumption 6 (Treatment monotonicity at the extensive margin). *The treatment monotonicity at the extensive margin assumption is given by*

$$\begin{aligned} Y_{i,t}^1 > 0 &\Rightarrow Y_{i,t}^0 > 0 \quad \forall i, \text{ or} \\ Y_{i,t}^0 > 0 &\Rightarrow Y_{i,t}^1 > 0 \quad \forall i. \end{aligned}$$

The assumption of *treatment monotonicity at the extensive margin* states that a positive outcome in case of treatment implies a positive outcome in case of no treatment or vice versa. Therefore, the treatment response is monotone with respect to the extensive margin decision. Given the potential outcome in case of treatment is positive, the potential outcome in case of no treatment is allowed to be higher or lower than the potential outcome in case of treatment. The assumption only requires that the potential outcome in case of no treatment is positive.

Assumption 7 (Time monotonicity at the extensive margin). *The time monotonicity at the extensive margin assumption is given by*

$$\begin{aligned} Y_{i,t}^0 > 0 &\Rightarrow Y_{i,t-1}^0 > 0 \quad \forall i, \text{ and} \\ Y_{i,t}^1 > 0 &\Rightarrow Y_{i,t-1}^1 > 0 \quad \forall i. \end{aligned}$$

The assumption of *time monotonicity at the extensive margin* states that a positive outcome in period t implies a positive outcome in period $t - 1$, both in case of treatment and no treatment. Thus, we assume that there are no individuals with a positive outcome in period t who have a zero outcome in period $t - 1$. The assumption rules out the possibility of time trends that affect the extensive margin decision. Given the potential outcome in period t is positive, the potential outcome in period $t - 1$ is allowed to be higher or lower than the potential outcome in period t . The assumption only requires that the potential outcome in period $t - 1$ is positive.

Proof. Assuming SUTVA, equation (16) can be rewritten to

$$\begin{aligned} \gamma_t^{DiD}(x) = & E(Y_{i,t}^1 - Y_{i,t-1}^1 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x) \\ & - E(Y_{i,t}^0 - Y_{i,t-1}^0 | Y_{i,t}^0 > 0, Y_{i,t-1}^0 > 0, D_i = 0, X_i = x). \end{aligned} \quad (17)$$

¹⁰ Using SUTVA, the *common trend in positive outcomes* assumption can be rewritten to $E(Y_{i,t}^0 - Y_{i,t-1}^0 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x) = E(Y_{i,t}^0 - Y_{i,t-1}^0 | Y_{i,t}^0 > 0, Y_{i,t-1}^0 > 0, D_i = 0, X_i = x)$, which is the version used in the proof of identification.

Adding and subtracting $E(Y_{i,t-1}^0 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x)$, and $E(Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x)$ to equation (17) and rearranging yields

$$\gamma_t^{DiD}(x) = E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x) \quad (18)$$

$$+ E(Y_{i,t-1}^0 - Y_{i,t-1}^1 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x) \quad (19)$$

$$+ E(Y_{i,t}^0 - Y_{i,t-1}^0 | Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0, D_i = 1, X_i = x) \quad (20)$$

$$+ E(Y_{i,t-1}^0 - Y_{i,t}^0 | Y_{i,t}^0 > 0, Y_{i,t-1}^0 > 0, D_i = 0, X_i = x). \quad (21)$$

Assuming *SUTVA* and *common trend in positive outcomes*, the sum of the two terms in (20) and (21) equals 0. Moreover, under the *no pre-treatment effect* assumption, the sum of the term in (19) is equal to zero. Assuming *time* and *treatment monotonicity at the extensive margin*, the term in (18) can be rewritten to $E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x)$.¹¹ This identifies the conditional-on- X version of the intensive margin average treatment effect on the treated.

The *common support* assumption then guarantees that all conditional-on- X versions of the IMATT exist. Based on (15), the conditional-on- X versions are aggregated with respect to the distribution of X in the subsample with $Y_{i,t}^1 > 0, Y_{i,t}^0 > 0$, and $D_i = 1$. Assuming *time* and *treatment monotonicity at the extensive margin*, this subsample is identical to the subsample with $Y_{i,t}^1 > 0, Y_{i,t-1}^1 > 0$, and $D_i = 1$.¹² By *SUTVA*, this subsample is again identical to the subsample with $Y_{i,t} > 0, Y_{i,t-1} > 0$, and $D_i = 1$, which is an observed subsample. $\square$

An obvious alternative to difference-in-difference is the simple difference estimator, given by

$$\gamma_t^D(x) = E(Y_{i,t} | Y_{i,t} > 0, D_i = 1, X_i = x) - E(Y_{i,t} | Y_{i,t} > 0, D_i = 0, X_i = x). \quad (22)$$

In Appendix A, we state sufficient conditions under which the simple difference estimator identifies the conditional-on- X intensive margin average treatment effect on the treated.

3.2 Special Case: Random Treatment

When treatment is randomly assigned, we do not need to condition on X to identify the causal effect. If we do not condition on X , we do not require the *common support* and the *no effect of treatment on covariates* assumptions. The other assumptions are still required to identify the intensive margin average treatment effect on the treated. A further implication of random treatment is that we can also identify the ATE, since the ATT equals the ATE under random treatment.

4 Estimation and Inference

Difference-in-difference estimation requires estimating different conditional expectations. Here we adopt a split sample approach. Let $\Delta Y_{i,t} = Y_{i,t} - Y_{i,t-1}$. For the difference-in-difference on positive outcomes estimator, we first estimate

$$m_1(x) = E(\Delta Y_{i,t} | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 1, X_i = x), \text{ and} \quad (23)$$

$$m_0(x) = E(\Delta Y_{i,t} | Y_{i,t} > 0, Y_{i,t-1} > 0, D_i = 0, X_i = x), \quad (24)$$

using ordinary least squares. That is, we regress $\Delta Y_{i,t}$ on X_i separately in the treated sample and in the untreated sample, restricted to the observations with positive outcomes in period t and $t - 1$. Since we condition

¹¹ By the *time monotonicity at the extensive margin*, the conditioning set can be reduced to $E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, D_i = 1, X_i = x)$. Using the *treatment monotonicity at the extensive margin*, the conditioning set can again be expanded to $E(Y_{i,t}^1 - Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x)$.

¹² Analogous to the rewriting of (18), see footnote 11.

the sample on observations with a positive outcome in period t and $t - 1$, we require panel data.¹³ Using the fitted functions $\widehat{m}_1(x)$ and $\widehat{m}_0(x)$, we then calculate fitted values $\widehat{m}_1(X_i)$ and $\widehat{m}_0(X_i)$. The intensive margin average treatment effect on the treated is then estimated as

$$\widehat{IMATT}_t^{DID} = \frac{1}{N_T} \sum_{\substack{i: Y_{i,t} > 0, \\ Y_{i,t-1} > 0, \\ D_i = 1}} [\widehat{m}_1(X_i) - \widehat{m}_0(X_i)], \quad (25)$$

where N_T is the number of treated observations with positive outcome in period t and $t - 1$.¹⁴

To conduct inference, we employ a nonparametric quantile bootstrap [15]. From the sample of observations with positive outcomes in period t and $t - 1$, we repeatedly draw a bootstrap sample of the same sample size. In the bootstrap sample, we estimate the IMATT as described above. This gives a distribution of bootstrap estimated IMATTs: $\widehat{IMATT}_t^1, \dots, \widehat{IMATT}_t^B$, where B is the number of bootstrap replications. We then construct a bootstrap estimated confidence interval as

$$[q_{\alpha/2}^*, q_{1-\alpha/2}^*], \quad (26)$$

where $q_{1-\alpha/2}^*$ is the $(1 - \alpha/2)$ -percentile of the distribution of bootstrap estimated IMATTs.

5 Empirical Application: Causal Effect of Reaching the Full Retirement Age on Working Hours

We apply the difference-in-difference methodology to estimate the causal intensive margin effect of reaching the full retirement age on working hours of women. We exploit a pension reform in Switzerland taking place in 2004. In this pension reform, the full retirement age (FRA) of women was increased from age 63 to age 64.¹⁵ This implies that women with year of birth 1941 or earlier reach FRA at age 63, while women with year of birth 1942 or later reach FRA at age 64. We use data from the Swiss Labor Force Survey (SLFS) from 2002-2009. The outcome of interest is working hours, denoted by $Y_{i,t}$.¹⁶ We restrict the sample to women aged 63. Therefore, treatment $D_i = 1$ for women who have reached FRA (year of birth 1941 or earlier), and $D_i = 0$ for women who have not reached FRA (year of birth 1942 or later). Since the reform affects individuals only based on their year of birth, assignment to treatment can be assumed to be almost random. We only include a categorical education variable and a dummy for being a Swiss citizen. We consider two estimation samples. The first sample consists of women with year of birth 1941 or 1942. That is, women exactly at the threshold of the pension reform. This sample is cleaner in terms of identification, but the number of observations decreases the power. For this reason we consider a second estimation sample, which includes women with year of birth 1939 to 1946. This sample includes more observations, but might pose a threat to identification if there is a time trend in working hours.

5.1 Discussion: Assumptions

With the exception of *time monotonicity at the extensive margin* and *common support*, we cannot directly test the identifying assumptions. Instead, we propose alternative tests that can be used to motivate the identifying

¹³ This marks a difference to the standard difference-in-difference estimator, for which it is also possible to use repeated cross-sections.

¹⁴ Alternatively, one could replace $\widehat{m}_1(X_i)$ in (25) with $\Delta Y_{i,t}$, such that it would not be necessary to estimate $\widehat{m}_1(x)$.

¹⁵ The FRA denotes the age at which a first pillar pension can be claimed without a deduction. There was a second pension reform in 2001, increasing the FRA of women from 62 to 63, which is not considered in this application.

¹⁶ We use contracted hours for wage employed and usual hours for self employed as our measure of working hours.

assumptions and discuss whether the assumptions are likely to be fulfilled in the context of our empirical application.¹⁷

SUTVA: This assumption cannot be tested. There is evidence for spillover effects within couples [16–18]. That is, the labor supply of one individual depends on whether the spouse has reached FRA. We are aware that this might pose a threat to identification, but assume that the spillover effects are negligible.

No pre-treatment effect: This assumption rules out that people adjust their working hours in anticipation of reaching FRA in the next period. We cannot directly test this assumption. We motivate the assumption by comparing the mean working hours in period $t - 1$, conditional on having positive working hours in period t and $t - 1$. The mean in the control group is 23.8 hours, in the treatment group 24.6 hours. A simple Welch two sample t-test does not reject the null hypothesis of equal means (p-value: 0.54). This indicates that the assumption is fulfilled. Moreover, if there is a pre-treatment effect, this effect will likely have the same sign as the treatment effect. As a result, the estimated treatment effect could be interpreted as a lower bound.

Common trend in positive outcomes: This assumption requires that the treatment group would experience the same time trend in working hours in case of no treatment as the control group. We cannot directly test this assumption, but we motivate the assumption by examining the pre-treatment trends of the control and treatment group. In Figure 1, we plot the mean working hours of women with positive hours in period t , $t - 1$ and $t - 2$. We observe that the trends between period $t - 2$ and $t - 1$ are roughly parallel, indicating that the assumption is fulfilled.

Figure 1: Assessment of Common Trend in Positive Outcomes

Note: Dots indicate the mean working hours of women aged 63 with positive working hours in period t , $t-1$ and $t-2$. Bars indicate the 95% normal approximation confidence interval for the mean. Year of birth between 1939 and 1946. Treated is the group which reaches FRA in period t (women with year of birth 1941 or earlier), Control is the group which does not reach FRA in period t (women with year of birth 1942 or later).

No effect of treatment on covariates: In the empirical application, we include a categorical education variable and a dummy for being a Swiss citizen. It seems unlikely that reaching FRA has an effect on these variables. If so, the effect is likely negligible.

Common support: This assumption can be tested. In each covariate cell, we calculate the fraction of treated observations. The results are presented in Table B1 in Appendix B. We observe that there is no covariate cell

¹⁷ The (alternative) tests are always based on the larger estimation sample, i.e. the sample including women with year of birth 1939 to 1946.

with only treated observations. Therefore, the assumption is fulfilled.

Treatment monotonicity at the extensive margin: This assumption rules out that people start to work because they reach FRA. There are indeed incentives to take up a job after reaching FRA. For example, part of the earnings are exempted from social security contributions. This increases the net wage. On the other hand, it seems plausible that reaching FRA either has no effect or drives people out of the labor market.

Time monotonicity at the extensive margin: This assumption can be tested. In the treated and control subsamples, we calculate the fraction of individuals with positive working hours in period t , conditional on not working in period $t - 1$. In the sample of women with year of birth 1939-46, 5% of the treated and 7.4% of the control sample state that they returned to work after having not worked in the period before. This poses a threat to our identification. However, the overall pattern in the age range 60-70 is that people rather leave the labor force as they become older.

5.2 Estimation Results

The results of the difference-in-difference estimation are presented in Table 2. In the estimation sample including only women with year of birth 1941-42 (left column), the estimated intensive margin average treatment effect on the treated is -5.003. That is, reaching FRA reduces the working hours of women with positive working hours irrespective of whether they have reached FRA or not on average by 5 hours. The bootstrap estimated 95% confidence interval does not include zero, indicating that the effect is statistically significantly different from zero. In the sample including women with year of birth 1939-46 (right column), the estimated intensive margin ATT is -4.215. Again, the bootstrap estimated 95% confidence interval does not include zero. This analysis provides evidence that women react at the intensive margin when reaching FRA.

Table 2: Results Difference-in-Difference on Positive Outcomes

	IMATT DID 1941-42	IMATT DID 1939-46
FRA reached	-5.003	-4.215
95% C.I.	[-8.47, -1.58]	[-6.58, -1.95]
Obs. (treat/cont)	63/87	156/405

Note: Confidence interval based on 1000 bootstrap replications. Sample includes women aged 63 with positive working hours in period t and $t-1$. Women with year of birth 1942 or later have FRA 64 (Control), women with year of birth 1941 or earlier have FRA 63 (Treated). The left column presents the results for women with year of the birth 1941-1942, and the right column those for women with year of birth 1939-1946.

6 Conclusion

This paper extends the literature on the identification and estimation of causal intensive margin effects. The intensive margin effect is of interest when subeffects are masked by the total effect. This is the case, for example, when the extensive and intensive margin effect have different signs. We use difference-in-difference methods to identify the causal intensive margin effect. We derive sufficient conditions under which the difference-in-difference estimator on positive outcomes identifies the causal intensive margin effect. We demonstrate that the difference-in-difference estimator on positive outcomes, compared to the standard difference-in-difference estimator, additionally requires *time* and *treatment monotonicity at the extensive margin*. We apply the methodology to estimate the causal intensive margin effect of reaching the full retirement age on working hours.

References

- [1] M.-J. Lee, Treatment Effects in Sample Selection Models and Their Nonparametric Estimation, *Journal of Econometrics*, vol. 167, no. 2, pp. 317–329, 2012.
- [2] K. Staub, A Causal Interpretation of Extensive and Intensive Margin Effects in Generalized Tobit Models, *The Review of Economics and Statistics*, vol. 96, no. 2, pp. 371–375, 2014.
- [3] M.-J. Lee, Extensive and Intensive Margin Effects in Sample Selection Models : Racial Effects on Wages, *Journal of the Royal Statistical Society. Series A: Statistics in Society*, vol. 180, pp. 817–839, 2017.
- [4] J. D. Angrist, Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors, *Journal of Business & Economic Statistics*, vol. 19, no. 1, pp. 2–28, 2001.
- [5] J. Tobin, Estimation of Relationships for Limited Dependent Variables, *Econometrica*, vol. 26, no. 1, pp. 24–36, 1958.
- [6] J. F. McDonald and R. A. Moffitt, The Uses of Tobit Analysis, *The Review of Economics and Statistics*, vol. 62, no. 2, pp. 318–321, 1980.
- [7] J. G. Cragg, Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods, *Econometrica*, vol. 39, no. 5, pp. 829–844, 1971.
- [8] N. Duan, W. G. Manning, C. N. Morris, and J. P. Newhouse, A Comparison of Alternative Models for the Demand for Medical Care, *Journal of Business & Economic Statistics*, vol. 1, no. 2, pp. 115–126, 1983.
- [9] J. J. Heckman, Sample Selection Bias as a Specification Error, *Econometrica*, vol. 47, no. 1, pp. 153–161, 1979.
- [10] C. E. Frangakis and D. B. Rubin, Principal Stratification in Causal Inference, *Biometrics*, vol. 58, no. 1, pp. 21–29, 2002.
- [11] J. D. Angrist, G. W. Imbens, and D. B. Rubin, Identification of Causal Effects Using Instrumental Variables, *Journal of the American Statistical Association*, vol. 91, pp. 444–455, 1996.
- [12] E. Deuchert, M. Huber, and M. Schelker, Direct and Indirect Effects Based on Difference-in-Differences With an Application to Political Preferences Following the Vietnam Draft Lottery, *Journal of Business & Economic Statistics*, vol. 37, no. 4, pp. 710–720, 2019.
- [13] M. Lechner, The Estimation of Causal Effects by Difference-in-Difference Methods, *Foundations and Trends in Econometrics*, vol. 4, no. 3, pp. 165–224, 2010.
- [14] D. B. Rubin, Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Educational Psychology*, vol. 66, no. 5, pp. 688–701, 1974.
- [15] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 1993.
- [16] J. Cribb, C. Emmerson, and G. Tetlow, Incentives, Shocks or Signals: Labour Supply Effects of Increasing the Female State Pension Age in the UK, IFS Working Paper, vol. W13/03, 2013.
- [17] E. Stanca, Couples' Retirement Under Individual Pension Design: A Regression Discontinuity Study for France, *Labour Economics*, vol. 49, pp. 14–26, 2017.
- [18] R. Lalive and P. Parrotta, How Does Pension Eligibility Affect Labor Supply in Couples?, *Labour Economics*, vol. 46, pp. 177–188, 2017.

A Identification Simple Difference Estimator on Positive Outcomes

The *simple difference estimator on positive outcomes* is given by

$$\gamma_t^D(x) = E(Y_{i,t} | Y_{i,t} > 0, D_i = 1, X_i = x) - E(Y_{i,t} | Y_{i,t} > 0, D_i = 0, X_i = x). \quad (\text{A27})$$

The following sufficient conditions identify the conditional-on- X intensive margin average treatment effect on the treated.

Proposition 2 (Identification simple difference estimator on positive outcomes). *Sufficient conditions to identify the causal intensive margin effect using the simple difference estimator on positive outcomes are:*

1. SUTVA (assumption 1),
2. no effect of treatment on covariates (assumption 4),
3. common support (assumption 5),
4. unconfoundedness (assumption 8), and
5. no switchers (assumption 9),

Or

5. conditional mean independence (assumption 10).

Assumption 8 (Unconfoundedness). *The unconfoundedness assumption is given by*

$$(Y_{i,t}^1, Y_{i,t}^0) \perp\!\!\!\perp D_i \mid X_i.$$

The *unconfoundedness* assumption requires that treatment is independent of the potential outcomes, conditional on covariates X_i .

Assumption 9 (No switchers). *The assumption of no switchers is given by*

$$Y_{i,t}^1 > 0 \Leftrightarrow Y_{i,t}^0 > 0 \quad \forall i.$$

The assumption of *no switchers* states that the potential outcome in case of treatment is positive if and only if the potential outcome in case of no treatment is positive. It therefore excludes the possibility that individuals have a positive outcome in case of treatment and a zero outcome in case of no treatment (switchers 1), or vice versa (switchers 2).

Assumption 10 (Conditional mean independence). *The conditional mean independence assumption is given by*

$$\begin{aligned} E(Y_{i,t}^1 | Y_{i,t}^1 > 0, Y_{i,t}^0 = 0, D_i = 1, X_i = x) &= E(Y_{i,t}^1 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x), \text{ and} \\ E(Y_{i,t}^0 | Y_{i,t}^1 = 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x) &= E(Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x). \end{aligned}$$

The assumption of *conditional mean independence* states that the expected potential outcome in case of treatment of switchers 1 is equal to the expected potential outcome in case of treatment of always-participants. Furthermore, the expected potential outcome in case of no treatment of switchers 2 is equal to the expected potential outcome in case of no treatment of always-participants.

Proof. Under *SUTVA* and *unconfoundedness*, and by the law of iterated expectations, equation (A27) can be rewritten as

$$\gamma_t^D(x) = \left[pE(Y_{i,t}^1 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x) \right. \quad (\text{A28})$$

$$\left. + (1 - p)E(Y_{i,t}^1 | Y_{i,t}^1 > 0, Y_{i,t}^0 = 0, D_i = 1, X_i = x) \right] \quad (\text{A29})$$

$$- \left[qE(Y_{i,t}^0 | Y_{i,t}^1 > 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x) \right. \quad (\text{A30})$$

$$\left. + (1 - q)E(Y_{i,t}^0 | Y_{i,t}^1 = 0, Y_{i,t}^0 > 0, D_i = 1, X_i = x) \right],$$

where $p \equiv P(Y_{i,t}^0 > 0 | Y_{i,t}^1 > 0, D_i = 1, X_i = x)$ and $q \equiv P(Y_{i,t}^1 > 0 | Y_{i,t}^0 > 0, D_i = 1, X_i = x)$. This term is equal to the causal intensive margin effect of interest in equation (14) if a) $p = q = 1$ (assumption of *no switchers*), or if b) the corresponding expectations in the brackets are identical, i.e. the expected potential outcome in case of treatment of switchers 1 is equal to the expected potential outcome of always-participants, and the expected potential outcome in case of no treatment of switchers 2 is equal to the expected potential outcome of always-participants (assumption of *conditional mean independence*). $\square$

B Common Support

Table B1: Analysis of Common Support

Secondary Education	Higher Education	Swiss citizen	Fraction Treated
0	0	0	0.091
0	0	1	0.308
0	1	0	0.350
0	1	1	0.262
1	0	0	0.154
1	0	1	0.298

Note: This table displays the fraction of treated observations (last column) for all possible covariate cells.