Identifying Dynamic LATEs with a Static Instrument*

Bruno Ferman† Otávio Tecchio‡

30th May 2023

Abstract

In many situations, researchers are interested in identifying dynamic effects of an irreversible treatment with a static binary instrumental variable (IV). For example, in evaluations of dynamic effects of training programs, with a single lottery determining eligibility. A common approach in these situations is to report per-period IV estimates. Under a dynamic extension of standard IV assumptions, we show that such IV estimators identify a weighted sum of treatment effects for different latent groups and treatment exposures. However, there is possibility of negative weights. We consider point and partial identification of dynamic treatment effects in this setting under different sets of assumptions.

Keywords: Instrumental Variables; Dynamic Local Average Treatment Effects; Negative Weights.

JEL Codes: C22; C23; C26.

*We would like to thank Vitor Possebom, Pedro Sant'Anna, Claudia Noack, Peter Hull, and seminar participants at FGV-EESP for helpful comments and suggestions.

†Sao Paulo School of Economics - FGV; email: bruno.ferman@fgv.br

‡Sao Paulo School of Economics - FGV; email: otavio.tecchio@fgv.br

1 Introduction

In many situations, researchers are interested in identifying dynamic effects of an irreversible treatment with a static binary instrumental variable (IV). As an example, consider evaluations of dynamic effects of training programs, exploiting a single lottery determining eligibility for a given cohort (e.g., Schochet, Burghardt and McConnell 2008; Alzúa, Cruces and Lopez 2016; Hirshleifer et al. 2016; Das 2021). Another example is the estimation of the dynamic effects of fertility on labor market outcomes using exogenous variations such as twins at first birth, sex composition of the first two children, and in-vitro fertilization success (e.g., Bronars and Grogger 1994; Angelov and Karimi 2012; Silles 2015; Lundborg, Plug and Rasmussen 2017). A common approach in these situations is to report per-period reduced form (RF) or IV estimates. However, we show that if units can access treatment at any period, these approaches may recover weighted sums of causal effects in which some weights are negative. In particular, we show that if first stages are decreasing over time, then there must be negative weights (and we may also have negative weights when the first stage is nondecreasing).

To understand the intuition behind the existence of negative weights, consider the example of a training program, where Z_i is a binary variable indicating the result of the lottery assignment for individual i. Consider individuals who, in the IV notation from Angrist, Imbens and Rubin (1996), are compliers in the first period. That is, if $Z_i = 1$, they attend the program, while if $Z_i = 0$, they do not attend the program. Now suppose some of those with $Z_i = 0$ end up attending the program in the second period. The RF estimator for the second period would compare the second-period average outcome between those with $Z_i = 1$ and $Z_i = 0$, in order to estimate the effect of the program one year after the assignment. However, the average outcome for those with $Z_i = 0$ would include those compliers in the first period that end up being treated in the second period. Therefore, the causal effect of being treated in the current period for these individuals enters negatively in the RF estimand. Since the counterpart of these individuals with $Z_i = 1$ were already treated in the previous period, those effects do not cancel out, unless treatment effects are independent of the time since treatment. We show that there also are other latent groups that generate problems in the per-period RF and IV comparisons.

We then consider the identification of the dynamic effects for first-period complies, which we call dynamic local average treatment effects (LATEs). We consider assumptions under which it is possible to solve for the dynamic LATEs recursively by correcting the reduced form estimands. In particular, our assumptions allow for unrestricted heterogeneity of causal effects with respect to treatment length, but instead assume homogeneity with respect to calendar time. As we discuss in more details in Section 2.2, such assumptions may be more palatable in some applications than assuming treatment effect homogeneity with respect to time since treatment. In such settings, our results indicate an alternative estimator relative to the standard per-period RF and IV estimators. The alternative we propose does not require additional sources of exogeneous variation (such as, for example, additional lotteries, or existence of excluded covariates).

Finally, we also consider partial identification of the dynamic LATEs, without requiring any assumption on the heterogeneous treatment effects. We do so by bounding the contamination term using bounds on the treatment effects. Bounds will be tighter the smaller the probability of late switching into treatment. This last result captures the intuition that contamination is less of a concern if a smaller proportion of individuals get treated after the first period. The idea of partially identifying treatment effects has been considered in other contexts by many others, for example, Manski (1990), Lee (2009), Flores and Flores-Lagunes (2013), and Bartalotti, Kédagni and Possebom (2023).

Our paper is related to a couple of papers that consider estimation of dynamic treatment effects in IV settings. Most related, Lundborg, Plug and Rasmussen (2017) recognize the shortcoming we study when considering per-period IV estimators in their estimation of dynamic effects of fertility on women's labor market outcomes. They present a simple model to provide an intuition on the direction of the bias in their setting in their Online Appendix C. Our Proposition 3.1 generalizes their arguments, with decomposition results for the RF and IV estimands in a more general IV setting with heterogeneous treatment effects and T periods. We also provide point and partial identification results in this setting. Miquel (2002) considers the identification of dynamic treatment effects with a static instrument. However, we consider sets of assumptions that are more reasonable for applications such as the estimation of dynamic effects of training programs or of fertility.[1] Han (2021) considers a very flexible IV framework, but his identification approach requires the existence of exogenous variables excluded from selection into treatment. In contrast, our identification results do not require the existence of covariates satisfying such assumptions. Heckman, Humphries and Veramendi (2016) consider a general discussion of the identification of dynamic treatment effects in dynamic discrete choice models. However, their model does not directly encompass the possibility of delayed treatment, which is the main source of contamination in our setting.

Our paper is also related to a literature on the estimation of dynamic treatment effects in alternative settings. For example, see Ding and Lehrer (2010), van den Berg and Vikström

[1] Miquel (2002) assumes that potential outcomes are independent of the instrument conditional on a history of treatment assignments. However, in the context of training programs or fertility, once we condition on a history of realized treatment assignment, we would be considering different latent groups, depending on whether $Z_i = 1$ or $Z_i = 0$.

(2022), and Bojinov, Rambachan and Shephard (2021). Their solutions, however, differ from ours in that they depend on controlling for fixed effects and lagged observables (Ding and Lehrer 2010), rely on sequential unconfoundedness assumptions (van den Berg and Vikström 2022), or rely on a sequence of randomizations (Bojinov, Rambachan and Shephard 2021). Most related to our solution for point identification, Cellini, Ferreira and Rothstein (2010) consider identification of dynamic effects in regression discontinuity designs using a recursive approach. However, they only consider the case of regression discontinuity designs that are sharp, and they focus on a different set of target parameters.

Our paper is also related to the literature on multiple treatment variables with lower dimensional instrumental variables (e.g., Torgovitsky 2015; D'Haultfœuille and Février 2015; Masten and Torgovitsky 2016; Caetano and Escanciano 2021; Hull 2018), since we can rewrite our setting as one with multiple treatments. However, we are able to exploit the particular dynamic structure of our setting to generate decomposition results for the per-period RF and IV estimands, and to propose alternative solutions relative to the ones in this literature.

Our work is also related to the literature on fuzzy difference-in-differences (de Chaisemartin and D'Haultfœuille 2017; Hudson, Hull and Liebersohn 2017). A crucial distinction from our setting is that we do not explore time variation under parallel trend assumptions, and we study two commonly used estimators: per-period reduced form and IV estimators. Moreover, we focus on the causal effect of exposure to treatment for multiple periods, while de Chaisemartin and D'Haultfœuille (2017) focus on the effects for switchers. Additionally, our identification results do not require the existence of a group for which the treatment rate is stable. Finally, our paper is also related to the recent papers on negative weights when considering two-way fixed effects estimators (de Chaisemartin and D'Haultfœuille 2020; Callaway and Sant'Anna 2021; Sun and Abraham 2021; Goodman-Bacon 2021; Athey and Imbens 2022; Borusyak, Jaravel and Spiess 2023), and when considering IV estimators with covariates (Kolesár 2013; Blandhol et al. 2022; Słoczyński 2022). However, the drivers of negative weights in our setting, and the solutions we propose, are different.

The remainder of this paper is organized as follows. In Section 2 we derive all of our main results for the simpler case in which we have only two periods. This includes decomposition results for the RF and IV estimands (Section 2.1), point identification results (Section 2.2), and partial identification results (Section 2.3). Then we consider the general multi-period setting in Section 3.

2 Two-periods setting

To simplify the exposition, consider first a setting with 2 periods of time, $t \in \{1, 2\}$. Units are indexed by i and time is indexed by t. We are interested in identifying dynamic effects of a binary treatment $D_{i,t}$ on some outcome $Y_{i,t}$. No unit is treated before the first period. There is selection into treatment, but we observe a static binary instrument Z_i.

We focus on settings where treatment is irreversible in the sense that once a unit is treated, it will be treated for all following periods. This is a common assumption in the difference-in-differences literature, and is known as staggered treatment adoption (e.g., Callaway and Sant'Anna 2021; Sun and Abraham 2021; Athey and Imbens 2022; Borusyak, Jaravel and Spiess 2023).

Assumption 2.1 (Irreversible Treatment). $D_{i,1} = 1 \implies D_{i,2} = 1$.

One advantage of focusing on irreversible treatments is that any possible sequence of treatment statuses at time t can be identified by zero if the unit has never been treated and by $(1, \tau)$ if the unit's first period of treatment was $t - \tau$. In this initial case with two periods, at $t = 1$ we may observe individuals with either treatment status 0 (not treated at $t = 1$) or $(1, 0)$ (treated at $t = 1$). In this case, $\tau = 0$ indicates that treatment length is zero, because the treatment started at $t = 1$, and we are considering the observation at $t = 1$. At $t = 2$, in addition to treatment status 0, we may have $(1, 1)$ (treated at $t = 1$, so $\tau = 1$ means that at $t = 2$ the length of the treatment is 1) or $(1, 0)$ (treated at $t = 2$).

We denote potential outcomes for unit i at time t by $Y_{i,t}(0, z)$ were i not treated at t and assigned by the instrument to z, and by $Y_{i,t}(1, \tau, z)$ were i first treated at $t - \tau$ and assigned by the instrument to z. Likewise, potential treatment statuses at period t are denoted by $D_{i,t}(z)$. We let AT_t denote always-takers at t (units such that $D_{i,t}(1) = D_{i,t}(0) = 1$), C_t denote compliers at t (units such that $D_{i,t}(1) > D_{i,t}(0)$), F_t denote defiers at t (units such that $D_{i,t}(1) < D_{i,t}(0)$) and NT_t denote never-takers at t (units such that $D_{i,t}(1) = D_{i,t}(0) = 0$).

In this setting with 2 periods, we could have, in principle, 16 latent groups, which are combinations of (AT_t, C_t, F_t, NT_t) for the two periods. Note, however, that Assumption 2.1 restricts these possibilities. In particular, the group AT_1 must also be AT_2. Moreover, the group C_1 must be either AT_2 (in case those with $Z_i = 0$ become treated in the second period) or C_2 (in case they remain untreated in the second period). In contrast, the group NT_1 can be any of the four possible latent groups in the second period even when treatment is irreversible. We say compliance is dynamic when there exists individuals whose latent groups change over time. Otherwise, we say that compliance is static. Compliance will be static if, for example, treatment can be accessed only in the first period.

For each $t \in \{1, 2\}$, define

$$RF_t := \mathbb{E}\left[Y_{i,t}|Z_i = 1\right] - \mathbb{E}\left[Y_{i,t}|Z_i = 0\right] \tag{1}$$

and

$$FS_t := \mathbb{E}\left[D_{i,t}|Z_i = 1\right] - \mathbb{E}\left[D_{i,t}|Z_i = 0\right], \tag{2}$$

the per-period reduced form and first stage estimands at t, respectively. Thus, whenever $FS_t \neq 0$, the per-period IV estimand at t is RF_t/FS_t.

As a first requirement for Z_i to be considered as an instrument, we propose a dynamic extension of the standard IV assumptions of Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996). The main difference from the assumptions in the static case is that we add independence and exclusion conditions in all periods.

Assumption 2.2. *Assume that the following hold:*

1. *Exclusion: For each $z \in \{0, 1\}$, $Y_{i,t}(0, z) = Y_{i,t}(0)$ and $Y_{i,t}(1, 0, z) = Y_{i,t}(1, 0)$ for $t \in \{1, 2\}$, and $Y_{i,2}(1, 1, z) = Y_{i,2}(1, 1)$.*

2. *Independence: $\left(Y_{i,1}(0), Y_{i,1}(1,0), Y_{i,2}(0), Y_{i,2}(1,0), Y_{i,2}(1,1), D_{i,1}(1), D_{i,1}(0), D_{i,2}(1), D_{i,2}(0)\right)$ is independent of Z_i.*

3. *Relevance at $t = 1$: $FS_1 \neq 0$.*

4. *Monotonicity at $t = 1$: $\mathbb{P}(F_1) = 0$.*

Our focus will be on comparisons between treated and untreated potential outcomes. Thus, our building blocks for decomposing the per-period reduced form estimands are causal effects of the form[2]

$$\Delta_t^\tau(g) := \mathbb{E}\left[Y_{it}(1, \tau) - Y_{it}(0)|g\right], \tag{3}$$

where g specifies a history of IV latent types. For example, an individual that is only treated in the first period if $Z_i = 1$ but, in the second period, gets treated regardless of Z_i belongs to $g = (C_1, AT_2)$. In this case, $\Delta_2^0(C_1, AT_2)$ is the treatment effect for this group of individuals at $t = 2$ when they received the treatment at $t = 2$. Note that there are two types of time heterogeneity in these treatment effects. The first one is with respect to the calendar time t and the second one is with respect to the treatment length τ.

We focus on target parameters of the type $\Delta_t^{t-1}(C_1)$, which we define as "dynamic LATEs". These are the local average treatment effects at time t, when treatment started

[2]Whenever written, expectations are assumed to exist.

at $t = 1$, for first-period compliers (C_1). For the comparison of effects across time to be valid, it is important that the IV latent type for which the causal effect is identified does not change. On the contrary, differences in effects across time cannot be solely attributed to time heterogeneity.

Given the notation above, it follows directly from Imbens and Angrist (1994) that $\Delta_1^0(C_1)$ is identified by the first period IV estimand under Assumption 2.2. Moreover, in case of static compliance, Assumptions 2.1 and 2.2 imply that the IV estimand in the second period identifies $\Delta_2^1(C_1)$, the effect at $t = 2$ of being treated at $t = 1$ for C_1 individuals. The argument for identification is analogous to the one for the first period estimand. For settings with T periods, the same is true for identification of $\Delta_t^{t-1}(C_1)$.

2.1 Decomposition of RF and IV estimands

While, under Assumptions 2.1 and 2.2, the IV estimands recover the dynamic LATEs when we have static compliance, we show that this would not generally be the case for $\Delta_2^1(C_1)$ when we have dynamic compliance.

Figure 1 depicts the remaining latent groups at $t = 2$ once we exclude first-period defiers (Assumption 2.2) and latent groups that are not consistent with irreversible treatment (Assumption 2.1). When we consider $RF_2 = \mathbb{E}[Y_{i,2}|Z_i = 1] - \mathbb{E}[Y_{i,2}|Z_i = 0]$, it is clear that the averages for $g = (AT_1, AT_2)$ cancel out, because the observed outcomes for those in this group are the same potential outcomes regardless of Z_i. The same is true for $g = (NT_1, AT_2)$ and $g = (NT_1, NT_2)$.

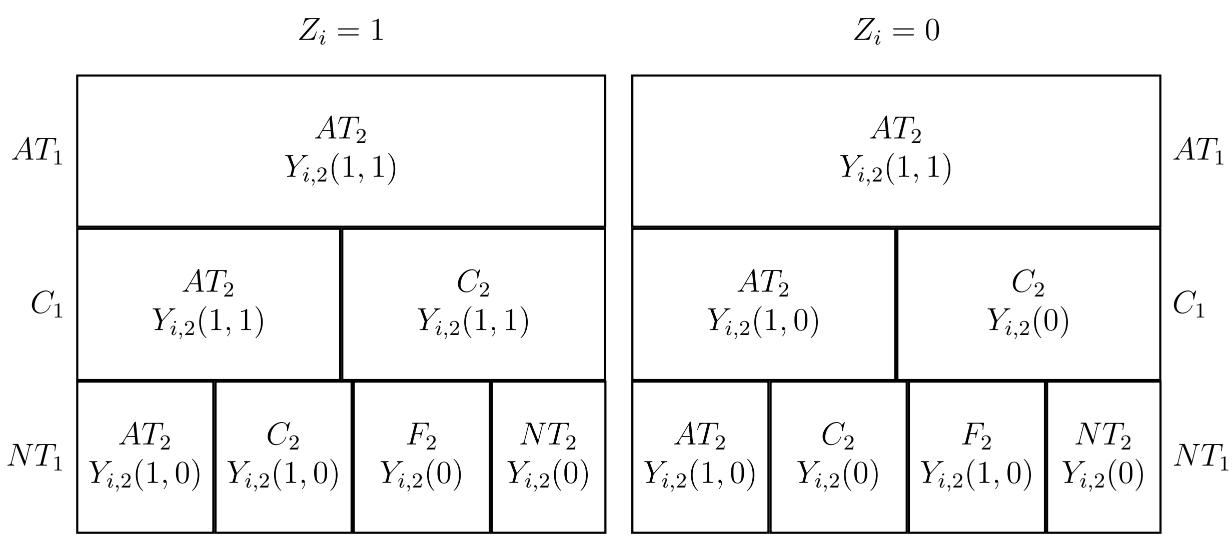

Figure 1: Latent groups and potential outcomes when $Z_i = 1$ and when $Z_i = 0$.

Therefore, RF_2 will capture the comparisons for the remaining latent groups. The main problem, however, is that for some of those groups the difference in observed outcomes between those with $Z_i = 1$ and $Z_i = 0$ does not represent a difference between potential outcomes $Y_{i,2}(1,1)$ and $Y_{i,2}(0)$. For example, we have that

$$\mathbb{E}[Y_{i,2}|Z_i = 1, C_1] - \mathbb{E}[Y_{i,2}|Z_i = 0, C_1] = \Delta_2^1(C_1) - \mathbb{P}(AT_2|C_1)\Delta_2^0(C_1, AT_2).$$

This will also be the case for the groups (NT_1, C_2), and (NT_1, F_2). The following proposition characterizes the RF_2 and FS_2 estimands in this setting with dynamic compliance.

Proposition 2.1. *Under Assumptions 2.1 and 2.2,*

$$\begin{aligned} RF_2 = &\mathbb{P}(C_1)\Delta_2^1(C_1) \\ &- \mathbb{P}(C_1, AT_2)\Delta_2^0(C_1, AT_2) - \mathbb{P}(NT_1, F_2)\Delta_2^0(NT_1, F_2) \\ &+ \mathbb{P}(NT_1, C_2)\Delta_2^0(NT_1, C_2) \end{aligned} \tag{4}$$

and

$$FS_2 = \mathbb{P}(C_1) - \mathbb{P}(C_1, AT_2) - \mathbb{P}(NT_1, F_2) + \mathbb{P}(NT_1, C_2). \tag{5}$$

Proof. Special case of Proposition 3.1. □

Equation 4 shows that the RF_2 depends on the dynamic LATE of interest at $t = 2$, $\Delta_2^1(C_1)$, but also on the effects for some groups that switch into treatment in the second period. In particular, because the (C_1, AT_2) and (NT_1, F_2) get treated at $t = 2$ only when $Z_i = 0$, the causal effect for them is negatively weighted. A negative weight for the (C_1, AT_2) is specially relevant because it implies that assuming no defiers in all periods is not sufficient to avoid negative weights overall. In fact, the decomposition for the FS_2 in Equation 5 shows that whenever $FS_2 < FS_1 = \mathbb{P}(C_1)$, there must be negative weights in RF_2 regardless of assumptions made on the existence of specific latent groups. More generally, for settings with T periods, we show in Corollary 3.1 that if there is a period in which the first stage is strictly smaller than in the period before, then there must be negative weights in the reduced form.

Equation 4 also indicates a typical case in which we may have sign reversal in the sense that the causal effects have all the same sign but the sign of RF_2 is opposite. Ignoring the NT_1's in RF_2 for the sake of the argument, if effects fade out sufficiently fast with respect to the treatment length dimension, then the term related to (C_1, AT_2) in RF_2 could be larger than the term related to C_1. For example, for the effects of children on parents' labor supply the treatment length dimension is the age of the child. Thus, if effects are always negative but decrease (in absolute value) when children get older, the reduced form estimand could be positive.

Once we have a decomposition for the reduced form and for the first stage, the decomposition for the IV estimand at $t = 2$ is immediate. Corollary 2.1 summarizes its main characteristics. The two main takeaways are that negative weights in RF_2 imply negative weights in the IV estimand and that the weights sum to one.

Corollary 2.1. *Under Assumptions 2.1 and 2.2, if $FS_2 \neq 0$, RF_2/FS_2 is a linear combination of the causal effects in Equation 4 in which the weights sum to one but some of them may be negative. There must be negative weights whenever $FS_2 < FS_1$. Moreover, the causal effects that are negatively weighted in RF_2/FS_2 are the same as in RF_2 if, and only if, $FS_2 > 0$.*

Proof. Special case of Corollary 3.1. □

Given the results above, we can consider the assumptions in which the second period IV estimand recovers $\Delta_2^1(C_1)$. We have already mentioned the case of static compliance. When compliance is static, individuals do not change treatment status from the first period to the second, implying

$$\mathbb{P}(C_1, AT_2) = \mathbb{P}(NT_1, C_2) = \mathbb{P}(NT_1, F_2) = 0,$$

and so RF_2 reduces to $\mathbb{P}(C_1)\Delta_2^1(C_1)$ while $FS_2 = \mathbb{P}(C_1)$. However, this is not the only case in which the IV estimand works. Assumption 2.3 formalizes types of treatment effects homogeneities which guarantee that the IV estimand at $t = 2$ identifies a causal effect.

Assumption 2.3. *For any latent group $g \in \{(C_1, AT_2), (NT_1, C_2), (NT_1, F_2)\}$ such that $\mathbb{P}(g) > 0$, $\Delta_2^1(C_1) = \Delta_2^0(g)$.*

Corollary 2.2. *Suppose Assumptions 2.1 and 2.2 hold. Under Assumption 2.3, and if $FS_2 \neq 0$,*

$$\frac{RF_2}{FS_2} = \Delta_2^1(C_1).$$

Proof. This result is immediate given Proposition 2.1. □

Assumption 2.3 is trivially satisfied if treatment effects are fully homogeneous (that is, with respect to treatment length, calendar time, and latent group). More generally, it says that, for those groups that contaminate the RF_2 estimand, their treatment effects at $t = 2$ must be the same as the LATE at $t = 2$ for the first-period compliers (who were treated at $t = 1$). This condition encompasses two sources of treatment effects homogeneity. First, it requires that treatment effects do not depend on the time since those individuals have been treated. This condition is arguably too strong in some of the settings we consider. For example, as already discussed, we should expect stronger effects of fertility on labor supply when the treatment length is smaller. Likewise, we may expect that training programs have negative effects in the beginning (while subjects are still taking classes), and then positive effects afterward. Second, Assumption 2.3 requires that those treatment effects (that do not depend on time since treatment) for the latent groups that contaminate the RF_2 are the same as for the first-period compliers, which is another kind of treatment effect homogeneity restriction. On the contrary, note that Assumption 2.3 does not impose restrictions on the possibility that treatment effects vary with calendar time.

Remark 1. *We note that defining potential outcomes as $\widetilde{Y}_{i,t}(1,z)$ when individual i is treated in the initial period, and $\widetilde{Y}_{i,t}(0,z)$ otherwise, would not be a valid solution without further assumptions. In this case, we would have that $\widetilde{Y}_{i,t}(0,z)$ would depend on z if we have dynamic compliance, so the usual IV exclusion restriction would not be valid for this definition of potential outcomes. For example, the instrument directly affects the potential outcome $\widetilde{Y}_{i,2}(0,z)$ for the (NT_1, C_2) individuals because they are treated at $t = 2$ only when $Z_i = 1$.*

2.2 Point identification of dynamic LATEs

Since the per-period IV estimands do not generally recover the dynamic LATEs, we consider alternative sets of assumptions in which these effects can be identified. In particular, we consider identification with unrestricted heterogeneity in causal effects with respect to the treatment length dimension. This comes at the cost of imposing homogeneity with respect to calendar time. In this sense, our approach is opposite to the per-period IV estimands in the type of time heterogeneity causal effects are allowed to have. Assumption 2.4 formalizes the types of homogeneity assumptions we require.

Assumption 2.4. *For any latent group $g \in \{(C_1, AT_2), (NT_1, C_2), (NT_1, F_2)\}$ such that $\mathbb{P}(g) > 0$, $\Delta_1^0(C_1) = \Delta_2^0(g)$.*

Assumption 2.4 says that, for those groups that contaminate RF_2, their average treatment effect at $t = 2$ must be the same as the first-period LATE. The main difference from Assumption 2.3 is that we have changed the type of time heterogeneity on causal effects. To understand the economic difference of these assumptions, it is useful to go back to the training program case. If, for example, the outcome of interest is employment, then we would expect the causal effect to vary whether the economy is in a recession or in a boom phase. Thus, homogeneity with respect to calendar time would be a strong assumption in a period of strong economic fluctuations. On the other hand, in periods of economic stability, it could be reasonable to assume that the effects do not depend on calendar time. Therefore, at least when we consider periods of economic stability, Assumption 2.4 should be more palatable than Assumption 2.3 in this type of applications.

We also note that the existence of latent groups (NT_1, C_2) and (NT_1, F_2) depends crucially on the empirical setting we are considering. Once more, consider the training program example. Suppose first that being lottery assigned to treatment implies that admission is guaranteed not only in the current period, but also in the following ones. In this case, we should expect that some of the NT_1 individuals would eventually get treatment in the second period if they have a guaranteed admission (in this case, if they have $Z_i = 1$). Therefore, we should expect $\mathbb{P}(NT_1, C_2) > 0$. Alternatively, suppose the lottery in the initial period does not guarantee admission in the following periods, and that never-takers from the first period do not receive different information depending on their Z_i. In this case, it would be more reasonable to assume that second-period take-up for NT_1 would not depend on instrument assignment, so $\mathbb{P}(NT_1, C_2) = \mathbb{P}(NT_1, F_2) = 0$. Therefore, in these settings, we would only require that $\Delta_1^0(C_1) = \Delta_2^0(C_1, AT_2)$. Other settings in which we only require this last homogeneity condition are the ones in which there are no NT_1 individuals. This would be the case when all individuals assigned to $Z_i = 1$ are treated in the first period.

Since $\Delta_1^0(C_1) = \Delta^0(C_1)$ is identified, under Assumption 2.4, it is possible to identify the contamination term of the reduced form estimand, and identify $\Delta_2^1(C_1)$ by correcting for the bias in RF_2.

Proposition 2.2. *Suppose Assumptions 2.1 and 2.2 hold. Under Assumption 2.4,*

$$\Delta^1(C_1) = \frac{RF_2}{FS_1} + \frac{\left(FS_1 - FS_2\right)}{FS_1}\frac{RF_1}{FS_1}. \tag{6}$$

Proof. Special case of Proposition 3.2. □

Therefore, Proposition 2.2 provides an alternative way to estimate dynamic LATEs that (relative to the per-period IV estimator) relies on arguably more reasonable assumptions in some settings. Moreover, note that, in contrast to the per-period IV estimand for $t = 2$, the identification result from Proposition 2.2 requires relevance of the instrument only in the first period (that is, we may have $FS_2 = 0$).

Remark 2. *Given the decomposition results from Proposition 2.1, it is possible to adapt the solution we propose in this section to other settings in which we have more information. For example, suppose we have another lottery at $t = 2$, that is independent from the first-period lottery, and let $\widetilde{C}_2$ be the compliers of this second lottery.*[3] *In this case, $\Delta_2^0(\widetilde{C}_2)$ is identified. Therefore, we may correct the contamination term using $\Delta_2^0(\widetilde{C}_2)$ (instead of*

[3] Individuals who participated in the first-period lottery may self select into participating in the second-period lottery. Moreover, lottery participants in this second-period lottery may also include individuals who did not participate in the first-period lottery.

$\Delta_1^0(C_1)$), assuming that, for any latent group $g \in \{(C_1, AT_2), (NT_1, C_2), (NT_1, F_2)\}$ such that $\mathbb{P}(g) > 0$, $\Delta_2^0(\widetilde{C}_2) = \Delta_2^0(g)$ (instead of Assumption 2.4). In this case, we can allow for unrestricted heterogeneity with respect to t and τ, but we still impose an homogeneity assumption on the treatment effects across some different latent groups.

Remark 3. *We can also consider extensions of our framework to analyze the causal effects of charter schools (Abdulkadiroğlu et al. 2011; Dobbie and Fryer 2011; Gleason et al. 2011; Angrist et al. 2016; Abdulkadiroğlu et al. 2016). For example, we can define potential outcomes $Y_{i,t}(s, \tilde{t})$ for a student i at time t, were he/she enrolled in a charter school for the first time at time $\tilde{t}$ in grade s. Then we can define causal effects based on comparisons between $Y_{i,t}(s, \tilde{t})$ and $Y_{i,t}(0)$, which is the potential outcome had the student never enrolled in a charter school until period t.[4] If we are considering a lottery at $t = 1$, we should take into account the possibility that students enroll in a charter school in subsequent periods, and our results can be adapted to this setting.*

2.3 Partial identification of dynamic LATEs

For settings in which the assumptions considered in Section 2.2 (or the assumptions that guarantee that per-period IV recovers the dynamic LATEs) would not be reasonable, we also consider partial identification results. In this case, we are able to construct bounds for the dynamic LATEs without imposing any assumption regarding the heterogeneity of the treatment effects. We derive bounds on the dynamic LATEs by considering bounds on the treatment effects. A particular case in which bounds for treatment effects are natural is settings with bounded outcomes (if there exist $\underline{Y}, \overline{Y} \in \mathbb{R}$ such that $\underline{Y} \leq Y_{i,2} \leq \overline{Y}$ with probability one, then the treatment effects are bounded, in absolute value, by $\overline{Y} - \underline{Y}$). We provide bounds that are valid without any assumption other than irreversible treatment (Assumption 2.1) and the basic conditions for IV validity (Assumption 2.2). We also show that it is possible to improve upon these general bounds by assuming that the treatment effects for the groups that contaminate RF_2 are homogeneous (given period and treatment length). For the sake of simplicity, Proposition 2.3 considers the case in which the lower

[4] Note that the way $Y_{i,t}(s, \tilde{t})$ is defined does not impose restrictions on the exposure to charter schools after this initial enrollment. In this case, we see the number of years enrolled in a charter school as one of the mechanisms in which the treatment (in this case, being enrolled in a charter school for the first time at time $\tilde{t}$ in grade s) may affect the outcomes. In the same way as college enrollment would be a mechanism in which charter school enrollment may affect earnings. An alternative in this case would be to define potential outcomes as a function of the number of years (or the specific years) in a charter school. Abdulkadiroğlu et al. (2016) present in their Appendix A the interpretation of the IV estimand when the treatment variable is given by the number of years enrolled in a charter school ($\tilde{d}$), and potential outcomes are defined as a function of $\tilde{d}$.

bound for the causal effects is nonpositive and the upper bound is nonnegative. In Appendix C we provide bounds without these restrictions.

Proposition 2.3. *Suppose Assumptions 2.1 and 2.2 hold. If there exist* $\underline{\Delta}, \overline{\Delta} \in \mathbb{R}$*, with* $\underline{\Delta} \leq 0 \leq \overline{\Delta}$*, such that for all* $g \in \{(C_1, AT_2), (NT_1, C_2), (NT_1, F_2)\}$ *with* $\mathbb{P}(g) > 0$*,* $\underline{\Delta} \leq \Delta_2^0(g) \leq \overline{\Delta}$*, then a lower bound for* $\Delta_2^1(C_1)$ *is given by*

$$\frac{RF_2}{FS_1} + \mathbb{P}\left(D_{i,2} > D_{i,1} | Z_i = 0\right) \frac{\underline{\Delta}}{FS_1} - \mathbb{P}\left(D_{i,2} > D_{i,1} | Z_i = 1\right) \frac{\overline{\Delta}}{FS_1} \tag{7}$$

and an upper bound is given by

$$\frac{RF_2}{FS_1} + \mathbb{P}\left(D_{i,2} > D_{i,1} | Z_i = 0\right) \frac{\overline{\Delta}}{FS_1} - \mathbb{P}\left(D_{i,2} > D_{i,1} | Z_i = 1\right) \frac{\underline{\Delta}}{FS_1}. \tag{8}$$

If, in addition to the conditions above, for all $g, g' \in \{(C_1, AT_2), (NT_1, C_2), (NT_1, F_2)\}$ *with* $\mathbb{P}(g) > 0$ *and* $\mathbb{P}(g') > 0$*,* $\Delta_2^0(g) = \Delta_0^2(g')$*, then*

$$\frac{RF_2}{FS_1} + \left[\mathbf{1}(FS_2 \leq FS_1)\underline{\Delta} + \mathbf{1}(FS_2 > FS_1)\overline{\Delta}\right] \frac{FS_1 - FS_2}{FS_1}, \tag{9}$$

where $\mathbf{1}(.)$ *is the indicator function, is a lower bound for* $\Delta_2^1(C_1)$ *and*

$$\frac{RF_2}{FS_1} + \left[\mathbf{1}(FS_2 \leq FS_1)\overline{\Delta} + \mathbf{1}(FS_2 > FS_1)\underline{\Delta}\right] \frac{FS_1 - FS_2}{FS_1} \tag{10}$$

is an upper bound. These bounds are (weakly) tighter than the previous ones.

Proof. Special case of Proposition 3.3. □

Remark 4. *Assuming* $\mathbb{P}(NT_1, C_2) = \mathbb{P}(NT_1, F_2) = 0$ *implies that the conditions in Proposition 2.3 for tighter bounds (Equations 9 and 10) hold. In Section 2.2, we discuss the settings in which assuming* $\mathbb{P}(NT_1, C_2) = \mathbb{P}(NT_1, F_2) = 0$ *should be reasonable. In this case, we would not need any assumption on the treatment effects heterogeneity to derive those tighter bounds. Moreover,* $\mathbb{P}(NT_1, C_2) = \mathbb{P}(NT_1, F_2) = 0$ *also implies* $FS_2 \leq FS_1$*, so that:*

$$\frac{RF_2}{FS_1} + \frac{FS_1 - FS_2}{FS_1}\underline{\Delta} \leq \Delta_2^1(C_1) \leq \frac{RF_2}{FS_1} + \frac{FS_1 - FS_2}{FS_1}\overline{\Delta}.$$

Remark 5. *An interesting case is when we consider the bounds in Equations 9 and 10 under a sign restriction for the treatment effects* $\Delta_2^0(g)$*. For example, if we assume causal effects are nonnegative* ($\underline{\Delta} = 0$)*, then* RF_2/FS_1 *would directly be the lower bound or upper bound* *(depending on whether FS_2 is lower than FS_1). In particular, if $FS_2 \leq FS_1$, RF_2/FS_1 is the lower bound.*

Importantly, the condition we consider to derive tighter bounds does not require any homogeneity assumption for the causal effects on the calendar period and treatment length dimensions. Note that for the more general bounds (Equations 7 and 8), the smaller the probability of late switching into treatment, the tighter the bounds. While for the bounds obtained in Equations 9 and 10, the smaller the changes in the first stage, the tighter the bounds.

3 T-periods setting

All of our results from Section 2 generalize for settings with an arbitrary number of periods. Consider a setting with T periods of time and let $\mathcal{T} \coloneqq \{1, ..., T\}$. The definitions of RF_t, FS_t, $D_{i,t}$, and latent groups extend naturally for this setting with T periods. We continue to assume that treatment is irreversible, so Assumption 2.1 becomes:

Assumption 3.1 (Irreversible Treatment). *For all $t \in \mathcal{T} \setminus \{T\}$, $D_{i,t} = 1 \implies D_{i,t+1} = 1$.*

Given irreversible treatment, we can again define the potential outcomes $Y_{i,t}(0, z)$, and $Y_{i,t}(1, \tau, z)$ depending only whether the unit has never been treated, or on whether it has been treated starting at period $t - \tau$. We also consider an extension of Assumption 2.2 for settings with T periods:

Assumption 3.2. *Assume that the following hold:*

1. *Exclusion: For each $t \in \mathcal{T}$ and $z \in \{0, 1\}$, $Y_{i,t}(0, z) = Y_{i,t}(0)$ and $Y_{i,t}(1, \tau, z) = Y_{i,t}(1, \tau)$ for all $\tau \in \{0, ..., t - 1\}$.*

2. *Independence: $\big(Y_{i,t}(0), Y_{i,t}(1, 0), ..., Y_{i,t}(1, t - 1), D_{i,1}(1), D_{i,1}(0), ..., D_{i,t}(1), D_{i,t}(0)\big)$ is independent of Z_i for all $t \in \mathcal{T}$.*

3. *Relevance at $t = 1$: $FS_1 \neq 0$.*

4. *Monotonicity at $t = 1$: $\mathbb{P}(F_1) = 0$.*

In this case, we are particularly interested in estimating the treatment effects $\Delta_t^{t-1}(C_1)$, which represent the local average treatment effects at time t of being treated $t - 1$ periods before (that is, when treatment started at $t = 1$), for the first period compliers. As before, the per-period IV estimand identifies $\Delta_t^{t-1}(C_1)$ under Assumption 3.2 if we have static compliance. However, this would not be the case when we have dynamic compliance.

3.1 Decomposition of RF and IV estimands with T periods

In order to generalize Proposition 2.1 for settings with T periods, we write $C_{t:t'}$ for units that are compliers from t to t', with analogous notation for defiers and never-takers. We only keep track of the first period in which units are always-takers because always-takers in a given period are always-takers in all following periods. Moreover, we define the following sets:

$$\begin{aligned}\mathcal{G}_2^+ &:= \big\{(NT_1, C_2)\big\},\\ \mathcal{G}_2^- &:= \big\{(C_1, AT_2), (NT_1, F_2)\big\},\end{aligned}$$

and, for each $t \in \mathcal{T} \setminus \{1, 2\}$,

$$\begin{aligned}\mathcal{G}_t^+ &:= \big\{(NT_{1:t-1}, C_t), (NT_{1:\ell-1}, F_{\ell:t-1}, AT_t) : \ell = 2, ..., t-1\big\},\\ \mathcal{G}_t^- &:= \big\{(C_{1:t-1}, AT_t), (NT_{1:t-1}, F_t), (NT_{1:\ell-1}, C_{\ell:t-1}, AT_t) : \ell = 2, ..., t-1\big\}.\end{aligned}$$

Assumption 3.1 implies that, for each $t \in \mathcal{T} \setminus \{1\}$, the latent groups in $\mathcal{G}_t^+$ are the ones that switch into treatment at t when $Z_i = 1$ and the latent groups in $\mathcal{G}_t^-$ are the ones that switch into treatment at t when $Z_i = 0$. Under this notation, we have the following decomposition results for the reduced forms and first stages.

Proposition 3.1. *Under Assumptions 3.1 and 3.2, for each $t \in \mathcal{T} \setminus \{1\}$,*

$$RF_t = \mathbb{P}(C_1)\, \Delta_t^{t-1}(C_1) - \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^-} \mathbb{P}(g)\, \Delta_t^{t-k}(g) + \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^+} \mathbb{P}(g)\, \Delta_t^{t-k}(g) \tag{11}$$

and

$$FS_t = \mathbb{P}(C_1) - \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^-} \mathbb{P}(g) + \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^+} \mathbb{P}(g). \tag{12}$$

Proof. See Appendix A.1. □

Corollary 3.1. *Under Assumptions 3.1 and 3.2, for any $t \in \mathcal{T} \setminus \{1\}$ such that $FS_t \neq 0$, RF_t/FS_t is a linear combination of the causal effects in Equation 11 in which the weights sum to one but some of them may be negative. A sufficient condition for the existence of negative weights at t is the existence of $k \in \{2, ..., t\}$ such that $FS_k < FS_{k-1}$. Moreover, the causal effects that are negatively weighted in RF_t/FS_t are the same as in RF_t if, and only if, $FS_t > 0$.*

Proof. See Appendix A.2. □

In Appendix B, we consider an alternative decomposition for the reduced form estimands. In particular, we show that they can be written as linear combinations of effects comparing treated and untreated potential outcomes, and effects comparing treated potential outcomes of different treatment lengths. The alternative decomposition helps to clarify in which situations we can simultaneously have the treatment making every unit better off and the reduced form being negative. This may be the case when treatment effects fade out with time since treatment.

3.2 Point identification with T periods

For the identification of dynamic LATEs, we consider again a recursive solution. For each $t \in \mathcal{T} \setminus \{1\}$, define

$$\mathcal{G}_t := \mathcal{G}_t^+ \cup \mathcal{G}_t^-,$$

the set of latent groups that switch into treatment at t and may contaminate the reduced form. The following assumption generalizes Assumption 2.4.

Assumption 3.3. *For all $t \in \mathcal{T}$ and $\tau \in \{0, ..., t-1\}$, $\Delta_t^\tau(C_1) = \Delta^\tau(C_1)$. Moreover, for each $t \in \mathcal{T} \setminus \{1\}$ and $\tau \in \{0, ..., t-2\}$, for any latent group $g \in \mathcal{G}_{t-\tau}$ such that $\mathbb{P}(g) > 0$, $\Delta^\tau(C_1) = \Delta_t^\tau(g)$.*

Proposition 3.2 below formalizes our identification result. To state it we consider matrix notation. Let $\mathbf{RF} := (RF_1, ..., RF_T)'$. For each $t \in \mathcal{T} \setminus \{1\}$, define $\rho_t := \mathbb{P}(D_{i,t} > D_{i,t-1} | Z_i = 0) - \mathbb{P}(D_{i,t} > D_{i,t-1} | Z_i = 1)$, the difference between the probability of switching into treatment for $Z_i = 0$ and $Z_i = 1$ units, which equals $FS_{t-1} - FS_t$ due to the irreversibility of treatment (Assumption 3.1). Moreover, let

$$\mathbf{P} := \begin{bmatrix} FS_1 & 0 & \dots & 0 \\ -\rho_2 & FS_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\rho_T & -\rho_{T-1} & \dots & FS_1 \end{bmatrix},$$

which is a lower triangular $T \times T$ matrix. Note that $\mathbf{P}$ is invertible provided that the instrument is relevant in the first period.

Proposition 3.2. *Suppose Assumptions 3.1 and 3.2 hold. Under Assumption 3.3*

$$\mathbf{\Delta} = \mathbf{P}^{-1}\mathbf{RF}, \tag{13}$$

where $\mathbf{\Delta} := \left(\Delta^0(C_1), ..., \Delta^{T-1}(C_1)\right)'$.

Proof. See Appendix A.3. □

3.3 Partial identification with T periods

In this general T-periods setting, we show that the dynamic LATEs are partially identified in every period for which the treatment effects are bounded (which, again, nests settings with bounded outcomes). As before, we provide bounds that do not require any homogeneity assumption on causal effects. We also show that with some homogeneities across latent groups, but with unrestricted heterogeneity on the treatment length and calendar time dimensions, we can obtain tighter bounds. Proposition 3.3 below generalizes Proposition 2.3. In Appendix C, we provide general bounds without requiring the lower bound (upper bound) for the treatment effects to be nonpositive (nonnegative).

Proposition 3.3. *Suppose Assumptions 3.1 and 3.2 hold. If, for $t \in \mathcal{T} \setminus \{1\}$, there exist $\underline{\Delta}_t, \overline{\Delta}_t \in \mathbb{R}$, with $\underline{\Delta}_t \leq 0 \leq \overline{\Delta}_t$, such that, for all $\tau \in \{0, ..., t-2\}$, if $g \in \mathcal{G}_{t-\tau}$ and $\mathbb{P}(g) > 0$, $\underline{\Delta}_t \leq \Delta_t^\tau(g) \leq \overline{\Delta}_t$, then a lower bound for $\Delta_t^{t-1}(C_1)$ is given by*

$$\frac{RF_t}{FS_1} + \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 0\right) \frac{\underline{\Delta}_t}{FS_1} - \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 1\right) \frac{\overline{\Delta}_t}{FS_1} \tag{14}$$

and an upper bound is given by

$$\frac{RF_t}{FS_1} + \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 0\right) \frac{\overline{\Delta}_t}{FS_1} - \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 1\right) \frac{\underline{\Delta}_t}{FS_1}. \tag{15}$$

If, in addition to the conditions above, for each $\tau \in \{0, ..., t-2\}$, for all $g, g' \in \mathcal{G}_{t-\tau}$ with $\mathbb{P}(g) > 0$ and $\mathbb{P}(g') > 0$, $\Delta_t^\tau(g) = \Delta_t^\tau(g')$, then

$$\frac{RF_t}{FS_1} + \underline{\Delta}_t \frac{(FS_1 - FS_t)}{FS_1} + \left(\overline{\Delta}_t - \underline{\Delta}_t\right) \sum_{k=2}^{t} \mathbf{1}\left(FS_{k-1} < FS_k\right) \frac{FS_{k-1} - FS_k}{FS_1} \tag{16}$$

is a lower bound for $\Delta_t^{t-1}(C_1)$ and

$$\frac{RF_t}{FS_1} + \overline{\Delta}_t \frac{(FS_1 - FS_t)}{FS_1} + \left(\underline{\Delta}_t - \overline{\Delta}_t\right) \sum_{k=2}^{t} \mathbf{1}\left(FS_{k-1} < FS_k\right) \frac{FS_{k-1} - FS_k}{FS_1} \tag{17}$$

is an upper bound for $\Delta_t^{t-1}(C_1)$. These bounds are (weakly) tighter than the previous ones.

Proof. See Appendix A.4. □

Remark 6. *The points in Remarks 4 and 5 generalize. Assuming that $\mathbb{P}(NT_{1:k-1}, C_k) = \mathbb{P}(NT_{1:k-1}, F_k) = 0$ for all $k \in \{2, ..., t\}$ implies that the conditions in Proposition 3.3 for* *tighter bounds hold at t and that first stages are nonincreasing (up to t). Under a sign restriction for treatment effects, if first stages are monotonic and the condition for tighter bounds holds, then RF_t/FS_1 is one of the bounds (whether it is the lower or upper bound depends on first stages being decreasing or increasing).*

References

Abdulkadiroğlu, Atila, Joshua Angrist, Susan Dynarski, Thomas J. Kane and Parag Pathak (2011). "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots". *The Quarterly Journal of Economics* 126.2, pp. 699–748.

Abdulkadiroğlu, Atila, Joshua D. Angrist, Peter D. Hull and Parag A. Pathak (2016). "Charters without Lotteries: Testing Takeovers in New Orleans and Boston". *American Economic Review* 106.7, pp. 1878–1920.

Alzúa, María Laura, Guillermo Cruces and Carolina Lopez (2016). "Long-Run Effects Of Youth Training Programs: Experimental Evidence From Argentina". *Economic Inquiry* 54.4, pp. 1839–1859.

Angelov, Nikolay and Arizo Karimi (2012). *Mothers' income recovery after childbearing.* Tech. rep. Working Paper.

Angrist, Joshua D., Sarah R. Cohodes, Susan M. Dynarski, Parag A. Pathak and Christopher R. Walters (2016). "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice". *Journal of Labor Economics* 34.2, pp. 275–318.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin (1996). "Identification of Causal Effects Using Instrumental Variables". *Journal of the American Statistical Association* 91.434, pp. 444–455.

Athey, Susan and Guido W. Imbens (2022). "Design-based analysis in Difference-In-Differences settings with staggered adoption". *Journal of Econometrics* 226.1. Annals Issue in Honor of Gary Chamberlain, pp. 62–79.

Bartalotti, Otávio, Désiré Kédagni and Vitor Possebom (2023). "Identifying marginal treatment effects in the presence of sample selection". *Journal of Econometrics* 234.2, pp. 565–584.

Blandhol, Christine, John Bonney, Magne Mogstad and Alexander Torgovitsky (2022). *When is TSLS Actually LATE?* Working Paper 29709. National Bureau of Economic Research.

Bojinov, Iavor, Ashesh Rambachan and Neil Shephard (2021). "Panel experiments and dynamic causal effects: A finite population perspective". *Quantitative Economics* 12.4, pp. 1171–1196.

Borusyak, Kirill, Xavier Jaravel and Jann Spiess (2023). "Revisiting Event Study Designs: Robust and Efficient Estimation". *arXiv preprint arXiv:2108.12419v3*.

Bronars, Stephen G. and Jeff Grogger (1994). "The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment". *The American Economic Review* 84.5, pp. 1141–1156.

Caetano, Carolina and Juan Carlos Escanciano (2021). “IDENTIFYING MULTIPLE MARGINAL EFFECTS WITH A SINGLE INSTRUMENT”. *Econometric Theory* 37.3, pp. 464–494.

Callaway, Brantly and Pedro H.C. Sant’Anna (2021). “Difference-in-Differences with multiple time periods”. *Journal of Econometrics* 225.2. Themed Issue: Treatment Effect 1, pp. 200–230.

Cellini, Stephanie Riegg, Fernando Ferreira and Jesse Rothstein (2010). “The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design”. *The Quarterly Journal of Economics* 125.1, pp. 215–261.

D’Haultfœuille, Xavier and Philippe Février (2015). “Identification of Nonseparable Triangular Models With Discrete Instruments”. *Econometrica* 83.3, pp. 1199–1210.

Das, Narayan (2021). “Training the disadvantaged youth and labor market outcomes: Evidence from Bangladesh”. *Journal of Development Economics* 149, p. 102585.

de Chaisemartin, Clément and Xavier D’Haultfœuille (2017). “Fuzzy Differences-in-Differences”. *The Review of Economic Studies* 85.2, pp. 999–1028.

— (2020). “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”. *American Economic Review* 110.9, pp. 2964–96.

Ding, Weili and Steven F Lehrer (2010). “Estimating Treatment Effects from Contaminated Multiperiod Education Experiments: The Dynamic Impacts of Class Size Reductions”. *The Review of Economics and Statistics* 92.1, pp. 31–42.

Dobbie, Will and Jr. Fryer Roland G. (2011). “Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children’s Zone”. *American Economic Journal: Applied Economics* 3.3, pp. 158–87.

Flores, Carlos A. and Alfonso Flores-Lagunes (2013). “Partial Identification of Local Average Treatment Effects With an Invalid Instrument”. *Journal of Business & Economic Statistics* 31.4, pp. 534–545.

Gleason, Philip, Melissa Clark, Christina Clark Tuttle and Emily Dwoyer (2011). *The Evaluation of Charter School Impacts*. Mathematica Policy Research Reports. Mathematica Policy Research.

Goodman-Bacon, Andrew (2021). “Difference-in-differences with variation in treatment timing”. *Journal of Econometrics* 225.2. Themed Issue: Treatment Effect 1, pp. 254–277.

Han, Sukjin (2021). “Identification in nonparametric models for dynamic treatment effects”. *Journal of Econometrics* 225.2. Themed Issue: Treatment Effect 1, pp. 132–147.

Heckman, James J., John Eric Humphries and Gregory Veramendi (2016). “Dynamic treatment effects”. *Journal of Econometrics* 191.2. Innovations in Measurement in Economics and Econometrics, pp. 276–292.

Hirshleifer, Sarojini, David McKenzie, Rita Almeida and Cristobal Ridao-Cano (2016). “The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey”. *The Economic Journal* 126.597, pp. 2115–2146.

Hudson, Sally, Peter Hull and Jack Liebersohn (2017). “Interpreting Instrumented Difference-in-Differences”. *Mimeo*.

Hull, Peter (2018). “IsoLATEing: Identifying Counterfactual-Specific Treatment Effects with Cross-Stratum Comparisons”. *Available at SSRN 2705108*.

Imbens, Guido W. and Joshua D. Angrist (1994). “Identification and Estimation of Local Average Treatment Effects”. *Econometrica* 62.2, pp. 467–475.

Kolesár, Michal (2013). *Estimation in an Instrumental Variables Model With Treatment Effect Heterogeneity*. Working Papers. Princeton University. Economics Department.

Lee, David S. (2009). “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects”. *The Review of Economic Studies* 76.3, pp. 1071–1102.

Lundborg, Petter, Erik Plug and Astrid Würtz Rasmussen (2017). “Can Women Have Children and a Career? IV Evidence from IVF Treatments”. *American Economic Review* 107.6, pp. 1611–37.

Manski, Charles F. (1990). “Nonparametric Bounds on Treatment Effects”. *The American Economic Review* 80.2, pp. 319–323.

Masten, Matthew A. and Alexander Torgovitsky (2016). “Identification of Instrumental Variable Correlated Random Coefficients Models”. *The Review of Economics and Statistics* 98.5, pp. 1001–1005.

Miquel, Ruth (2002). “Identification of Dynamic Treatment Effects by Instrumental Variables”. *Mimeo*.

Schochet, Peter Z., John Burghardt and Sheena McConnell (2008). “Does Job Corps Work? Impact Findings from the National Job Corps Study”. *American Economic Review* 98.5, pp. 1864–86.

Silles, Mary A. (2015). “The impact of children on women’s labour supply and earnings in the UK: evidence using twin births”. *Oxford Economic Papers* 68.1, pp. 197–216.

Słoczyński, Tymon (2022). “When Should We (Not) Interpret Linear IV Estimands as LATE?” *arXiv preprint arXiv:2011.06695v6*.

Sun, Liyang and Sarah Abraham (2021). “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects”. *Journal of Econometrics* 225.2. Themed Issue: Treatment Effect 1, pp. 175–199.

Torgovitsky, Alexander (2015). “Identification of Nonseparable Models Using Instruments With Small Support”. *Econometrica* 83.3, pp. 1185–1197.

van den Berg, Gerard J. and Johan Vikström (2022). "Long-Run Effects of Dynamically Assigned Treatments: A New Methodology and an Evaluation of Training Effects on Earnings". *Econometrica* 90.3, pp. 1337–1354.

A Proofs

A.1 Proof of Proposition 3.1

Fix $t \in \mathcal{T} \setminus \{1\}$. Under Assumption 3.2, the only latent groups that do not have equal potential outcomes (in expectation) when assigned to different instrument values are the ones that would behave differently if assigned to $Z_i = 1$ or $Z_i = 0$. Thus, units that are always-takers in all periods, units that are never-takers in all periods up to t, and units such that $(NT_{1:k-1}, AT_k)$ for some $k \in \{2, ..., t\}$ do not show up in our decomposition. The terms related to them cancel out.

Assumptions 3.1 and 3.2 imply that C_1, $(NT_{1:k-1}, C_k)$ or $(NT_{1:k-1}, F_k)$ with $k \in \{2, ..., t\}$ are the only groups that can have different potential treatment status depending on Z_i at t. Moreover, at each $k \in \{2, ..., t\}$, NT_{k-1} units' behavior parallels the behavior of all units in the first period, except that we allow for defiance. In particular, because of Assumption 3.1, treatment access for $(NT_{1:k-1}, C_k)$ and $(NT_{1:k-1}, F_k)$ groups, with $k \in \{2, ..., t\}$, has a dynamic that is analogous to the one for the C_1 group. Therefore, it suffices to consider the decomposition of $\mathbb{E}[Y_{i,t}|Z_i = 1, C_1] - \mathbb{E}[Y_{i,t}|Z_i = 0, C_1]$. Decomposition of the other terms follows from similar calculations, noting that defiers enter RF_t with opposite signs.

From Assumption 3.1, C_1 units with $Z_i = 1$ are treated in all periods and so

$$\mathbb{E}[Y_{i,t}|Z_i = 1, C_1] = \mathbb{E}[Y_{i,t}(1, t-1)|C_1] \tag{18}$$

follows from Assumption 3.2. To relate $\mathbb{E}[Y_{i,t}|Z_i = 0, C_1]$ to potential outcomes, we need to consider all possible latent group histories C_1 units can take up to t. Under Assumption 3.1, these histories have the form $(C_{1:k-1}, AT_k)$ with $k \in \{2, ..., t\}$ or $C_{1:t}$. Working forwardly and applying Assumption 3.2, we get:

$$\begin{aligned}
\mathbb{E}\left[Y_{i,t}|Z_i = 0, C_1\right] &= \mathbb{P}\left(AT_2|C_1\right)\mathbb{E}\left[Y_{i,t}(1, t-2)|C_1, AT_2\right] \\
&\quad + \mathbb{P}\left(C_2|C_1\right)\mathbb{E}\left[Y_{i,t}|Z_i = 0, C_{1:2}\right] \\
&= \mathbb{P}\left(AT_2|C_1\right)\mathbb{E}\left[Y_{i,t}(1, t-2)|C_1, AT_2\right] \\
&\quad + \mathbb{P}\left(C_2|C_1\right)\Big\{\mathbb{P}\left(AT_3|C_{1:2}\right)\mathbb{E}\left[Y_{i,t}(1, t-3)|C_{1:2}, AT_3\right] \\
&\qquad + \mathbb{P}\left(C_3|C_{1:2}\right)\Big[\mathbb{P}\left(AT_4|C_{1:3}\right)\mathbb{E}\left[Y_{i,t}(1, t-4)|C_{1:3}, AT_4\right] \\
&\qquad\quad + ...\mathbb{P}\left(C_{t-1}|C_{1:t-2}\right)\Big(\mathbb{P}\left(AT_t|C_{1:t-1}\right)\mathbb{E}\left[Y_{i,t}(1, 0)|C_{1:t-1}, AT_t\right] \\
&\qquad\qquad + \mathbb{P}\left(C_t|C_{1:t-1}\right)\mathbb{E}\left[Y_{i,t}(0)|C_{1:t}\right]\Big)...\Big]\Big\}.
\end{aligned} \tag{19}$$

Noting that $\mathbb{E}\left[Y_{i,t}(1,0)|C_{1:t-1}, AT_t\right] = \mathbb{E}\left[Y_{i,t}(0)|C_{1:t-1}, AT_t\right] + \Delta_t^0(C_{1:t-1}, AT_t)$, it follows from the Law of Iterated Expectations that the last term in parenthesis in the expression for $\mathbb{E}[Y_{i,t}|Z_i = 0, C_1]$ equals $\mathbb{P}\left(AT_t|C_{1:t-1}\right)\Delta_t^0(C_{1:t-1}, AT_t) + \mathbb{E}\left[Y_{i,t}(0)|C_{1:t-1}\right]$.

Repeating this process backwards, we obtain:

$$\mathbb{E}\left[Y_{i,t}|Z_i = 0, C_1\right] = \mathbb{E}\left[Y_{i,t}(0)|C_1\right] + \sum_{k=2}^{t}\left(\prod_{\ell=2}^{k-1}\mathbb{P}\left(C_\ell|C_{1:\ell-1}\right)\right)\mathbb{P}\left(AT_k|C_{1:k-1}\right)\Delta_t^{t-k}\left(C_{1:k-1}, AT_k\right),$$

under the convention that $\prod_{\ell=2}^{1} ... = 1$. Lastly, write the product of probabilities as a joint probability to get:

$$\mathbb{E}\left[Y_{i,t}|Z_i = 0, C_1\right] = \mathbb{E}\left[Y_{i,t}(0)|C_1\right] + \sum_{k=2}^{t}\mathbb{P}\left(C_{1:k-1}, AT_k|C_1\right)\Delta_t^{t-k}\left(C_{1:k-1}, AT_k\right),$$

which implies:

$$\mathbb{E}\left[Y_{i,t}|Z_i = 1, C_1\right] - \mathbb{E}\left[Y_{i,t}|Z_i = 0, C_1\right] = \Delta_t^{t-1}(C_1) - \sum_{k=2}^{t}\mathbb{P}\left(C_{1:k-1}, AT_k|C_1\right)\Delta_t^{t-k}\left(C_{1:k-1}, AT_k\right).$$

Computing the equivalent decomposition for each of the other histories and accounting for the probability of each of them, we get:

$$\begin{aligned}
RF_t = {} & \mathbb{P}\left(C_1\right)\Delta_t^{t-1}(C_1) \\
& - \sum_{k=2}^{t}\mathbb{P}\left(C_{1:k-1}, AT_k\right)\Delta_t^{t-k}\left(C_{1:k-1}, AT_k\right) \\
& + \sum_{k=2}^{t}\Bigg[\mathbb{P}\left(NT_{1:k-1}, C_k\right)\Delta_t^{t-k}\left(NT_{1:k-1}, C_k\right) \\
& \qquad - \sum_{\ell=k+1}^{t}\mathbb{P}\left(NT_{1:k-1}, C_{k:\ell-1}, AT_\ell\right)\Delta_t^{t-\ell}\left(NT_{1:k-1}, C_{k:\ell-1}, AT_\ell\right)\Bigg] \\
& - \sum_{k=2}^{t}\Bigg[\mathbb{P}\left(NT_{1:k-1}, F_k\right)\Delta_t^{t-k}\left(NT_{1:k-1}, F_k\right) \\
& \qquad - \sum_{\ell=k+1}^{t}\mathbb{P}\left(NT_{1:k-1}, F_{k:\ell-1}, AT_\ell\right)\Delta_t^{t-\ell}\left(NT_{1:k-1}, F_{k:\ell-1}, AT_\ell\right)\Bigg],
\end{aligned} \tag{20}$$

under the convention that $\sum_{\ell=t+1}^{t} ... = 0$. Note that $\sum_{k=2}^{t}\sum_{\ell=k+1}^{t} ...$ under the convention $\sum_{\ell=t+1}^{t} ...0$ can be written as $\sum_{\ell=2}^{t}\sum_{k=2}^{\ell-1} ...$ under the convention $\sum_{k=2}^{1} ... = 0$. Thus, rearranging Equation 20 and changing the index in the double sums (so that the outer summation is indexed by k and the inner one by ℓ with appropriate adjustment in the subscripts), we obtain:

$$
\begin{aligned}
RF_t =\ & \mathbb{P}\left(C_1\right) \Delta_t^{t-1}(C_1) - \sum_{k=2}^{t} \mathbb{P}\left(C_{1:k-1}, AT_k\right) \Delta_t^{t-k}\left(C_{1:k-1}, AT_k\right) \\
& + \sum_{k=2}^{t} \Bigg[\mathbb{P}\left(NT_{1:k-1}, C_k\right) \Delta_t^{t-k}\left(NT_{1:k-1}, C_k\right) - \mathbb{P}\left(NT_{1:k-1}, F_k\right) \Delta_t^{t-k}\left(NT_{1:k-1}, F_k\right) \Bigg] \\
& - \sum_{k=2}^{t} \sum_{\ell=2}^{k-1} \Bigg[\mathbb{P}\left(NT_{1:\ell-1}, C_{\ell:k-1}, AT_k\right) \Delta_t^{t-k}\left(NT_{1:\ell-1}, C_{\ell:k-1}, AT_k\right) \\
& \qquad - \mathbb{P}\left(NT_{1:\ell-1}, F_{\ell:k-1}, AT_k\right) \Delta_t^{t-k}\left(NT_{1:\ell-1}, F_{\ell:k-1}, AT_k\right) \Bigg].
\end{aligned}
\tag{21}
$$

The result as stated in Equation 11 follows from noting that for each $k \in \{2, ..., t\}$, any group g for which the causal effect $\Delta_t^{t-k}(g)$ appears in Equation 21 multiplied by a negative (respectively, positive) probability is such that $g \in \mathcal{G}_k^-$ (respectively, $g \in \mathcal{G}_k^+$).

For FS_t, we get from an analogous argument:

$$
\begin{aligned}
FS_t = \mathbb{P}\left(C_1\right) - \sum_{k=2}^{t} \mathbb{P}\left(C_{1:k-1}, AT_k\right) + \sum_{k=2}^{t} \Bigg[\mathbb{P}\left(NT_{1:k-1}, C_k\right) - \mathbb{P}\left(NT_{1:k-1}, F_k\right) \Bigg] \\
- \sum_{k=2}^{t} \sum_{\ell=2}^{k-1} \Bigg[\mathbb{P}\left(NT_{1:\ell-1}, C_{\ell:k-1}, AT_k\right) - \mathbb{P}\left(NT_{1:\ell-1}, F_{\ell:k-1}, AT_k\right) \Bigg],
\end{aligned}
\tag{22}
$$

under the convention that $\sum_{\ell=2}^{1} ... = 0$. Again, the result as stated in Equation 12 follows from the definition of the sets $\mathcal{G}_k^-$'s and $\mathcal{G}_k^+$'s.

A.2 Proof of Corollary 3.1

That RF_t/FS_t is a linear combination of the causal effects in RF_t is straightforward. That the weights in the IV estimand sum to one follows from noting that the sum of the probabilities in RF_t equals FS_t. For any given $k \in \{2, ..., t\}$, we have that

$$
FS_k - FS_{k-1} = - \sum_{g \in \mathcal{G}_k^-} \mathbb{P}(g) + \sum_{g \in \mathcal{G}_k^+} \mathbb{P}(g) < 0
$$

only when there exists $g \in \mathcal{G}_k^-$ such that $\mathbb{P}(g) > 0$, which implies that there is at least one causal effect that enters RF_t multiplied by a negative probability, which in turn implies a negative weight in the IV estimand at t. Lastly, $FS_t > 0$ is a necessary and sufficient condition for the negatively weighted causal effects in RF_t and RF_t/FS_t to be the same because the sign of the weights in the IV estimand equals the sign of the weights in RF_t times the sign of FS_t.

A.3 Proof of Proposition 3.2

For any $t \in \mathcal{T} \setminus \{1\}$,

$$\rho_t = FS_{t-1} - FS_t = \sum_{g \in \mathcal{G}_t^-} \mathbb{P}(g) - \sum_{g \in \mathcal{G}_t^+} \mathbb{P}(g).$$

Under Assumption 3.3, for any given $t \in \mathcal{T} \setminus \{1\}$, RF_t becomes

$$\begin{aligned} RF_t &= \mathbb{P}\left(C_1\right) \Delta^{t-1}(C_1) - \left[\sum_{k=2}^{t} \left(\sum_{g \in \mathcal{G}_k^-} \mathbb{P}\left(g\right) - \sum_{g \in \mathcal{G}_k^+} \mathbb{P}\left(g\right) \right) \Delta^{t-k}\left(C_1\right) \right] \\ &= \mathbb{P}\left(C_1\right) \Delta^{t-1}(C_1) - \sum_{k=2}^{t} \rho_k \Delta^{t-k}\left(C_1\right), \end{aligned}$$

which implies the linear system $\mathbf{RF} = \mathbf{P\Delta}$ if we recall that $FS_1 = \mathbb{P}(C_1)$ and that $RF_1 = \mathbb{P}(C_1)\Delta^0(C_1)$ under Assumption 3.3. The desired result follows from $\mathbf{P}$ being invertible under Assumption 3.2.

A.4 Proof of Proposition 3.3

For the bounds that are valid only assuming 3.1 and 3.2, we prove the more general version (as stated in Appendix C). Fix $t \in \mathcal{T} \setminus \{1\}$. Rearranging the reduced form (Equation 11):

$$\begin{aligned} \mathbb{P}\left(C_1\right) \Delta_t^{t-1}(C_1) &= RF_t + \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^-} \mathbb{P}\left(g\right) \Delta_t^{t-k}\left(g\right) - \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^+} \mathbb{P}\left(g\right) \Delta_t^{t-k}\left(g\right) \\ &\geq RF_t + \underline{\Delta}_t \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^-} \mathbb{P}\left(g\right) - \overline{\Delta}_t \sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^+} \mathbb{P}\left(g\right) \end{aligned} \quad (23)$$

Notice that $\sum_{k=2}^{t} \sum_{g \in \mathcal{G}_k^-} \mathbb{P}\left(g\right) \leq \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 0\right)$ because there are latent groups that switch into treatment after the first period when $Z_i = 0$ that are not included in the sets $\mathcal{G}_k^-$ for any $k \in \{2, ..., t\}$ (namely, the $(NT_{1:k-1}, AT_k)$ with $k \in \{2, ..., t\}$). Moreover, $\sum_{k=2}^t \sum_{g \in \mathcal{G}_k^-} \mathbb{P}(g) \geq \max\{FS_1 - FS_t, 0\}$. Also, $\max\{FS_t - FS_1, 0\} \leq \sum_{k=2}^t \sum_{g \in \mathcal{G}_k^+} \mathbb{P}(g) \leq \mathbb{P}(D_{i,t} > D_{i,1} | Z_i = 1)$. Thus, we can get a lower bound for the expression in the second row of Equation 23 by bounding the sum of probabilities, which implies the the following lower bound for $\mathbb{P}(C_1)\, \Delta_t^{t-1}(C_1)$:

$$\begin{aligned} &RF_t + \mathbf{1}\left(\underline{\Delta}_t < 0\right) \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 0\right) \underline{\Delta}_t + \mathbf{1}\left(\underline{\Delta}_t \geq 0\right) \max\left\{FS_1 - FS_t, 0\right\} \underline{\Delta}_t \\ &\quad - \mathbf{1}\left(\overline{\Delta}_t \geq 0\right) \mathbb{P}\left(D_{i,t} > D_{i,1} | Z_i = 1\right) \overline{\Delta}_t - \mathbf{1}\left(\overline{\Delta}_t < 0\right) \max\left\{FS_t - FS_1, 0\right\} \overline{\Delta}_t, \end{aligned}$$

from which the lower bound in Equation 14 follows directly since $\mathbb{P}(C_1) = FS_1 > 0$ under Assumption 3.2. The argument for the upper bound is analogous.

To obtain the bounds under the condition that for each $\tau \in \{0, ..., t-2\}$, for all $g, g' \in \mathcal{G}_{t-\tau}$ with $\mathbb{P}(g) > 0$ and $\mathbb{P}(g') > 0$, $\Delta_t^\tau(g) = \Delta_t^\tau(g')$, note that under such condition RF_t Equation (11) becomes

$$\mathbb{P}(C_1)\, \Delta_t^{t-1}(C_1) = RF_t + \sum_{k=2}^t \rho_k \Delta_t^{t-k}(*),$$

where, for a given $k \in \{2, ..., t\}$, $\Delta_t^{t-k}(*) \in \left[\underline{\Delta}_t, \overline{\Delta}_t\right]$ equals $\Delta_t^{t-k}(g)$ for all $g \in \mathcal{G}_k$. Then, because for any $k \in \{2, ..., t\}$, $\rho_k = FS_{k-1} - FS_k$,

$$\begin{aligned} \mathbb{P}(C_1)\, \Delta_t^{t-1}(C_1) &\geq RF_t + \sum_{k=2}^t \mathbf{1}\left(FS_{k-1} \geq FS_k\right) \rho_k \underline{\Delta}_t + \sum_{k=2}^t \mathbf{1}\left(FS_{k-1} < FS_k\right) \rho_k \overline{\Delta}_t \\ &= RF_t + \underline{\Delta}_t \sum_{k=2}^t \rho_k + \left(\overline{\Delta}_t - \underline{\Delta}_t\right) \sum_{k=2}^t \mathbf{1}\left(FS_{k-1} < FS_k\right) \rho_k \end{aligned} \tag{24}$$

and the upper bound follows from an analogous argument. The bounds as stated in the proposition follow from $\sum_{k=2}^t \rho_k = FS_1 - FS_t$. To prove that these later bounds are tighter, we note, from comparing Equations 23 and 24, that a sufficient condition for the lower bound to be tighter is

$$\begin{aligned} &\left[\sum_{k=2}^t \mathbf{1}\left(FS_{k-1} \geq FS_k\right) \rho_k - \sum_{k=2}^t \sum_{g \in \mathcal{G}_k^-} \mathbb{P}(g)\right] \underline{\Delta}_t \\ + &\left[\sum_{k=2}^t \mathbf{1}\left(FS_{k-1} < FS_k\right) \rho_k + \sum_{k=2}^t \sum_{g \in \mathcal{G}_k^+} \mathbb{P}(g)\right] \overline{\Delta}_t \geq 0, \end{aligned}$$

which is equivalent to

$$\begin{aligned}
&\left[\sum_{k=2}^{t}\rho_k-\sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^-}\mathbb{P}\left(g\right)+\sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^+}\mathbb{P}\left(g\right)\right]\underline{\Delta}_t\\
&+\left[\sum_{k=2}^{t}\mathbf{1}\left(FS_{k-1}<FS_k\right)\rho_k+\sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^+}\mathbb{P}\left(g\right)\right]\left(\overline{\Delta}_t-\underline{\Delta}_t\right)\geq 0\\
\iff &\left[\sum_{k=2}^{t}\mathbf{1}\left(FS_{k-1}<FS_k\right)\rho_k+\sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^+}\mathbb{P}\left(g\right)\right]\left(\overline{\Delta}_t-\underline{\Delta}_t\right)\geq 0
\end{aligned}$$

since $\sum_{k=2}^{t}\rho_k-\sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^-}\mathbb{P}\left(g\right)+\sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^+}\mathbb{P}\left(g\right)=0$. Because $\overline{\Delta}_t-\underline{\Delta}_t\geq 0$ and

$$\begin{aligned}
-\sum_{k=2}^{t}\mathbf{1}\left(FS_{k-1}<FS_k\right)\rho_k &= \sum_{k=2}^{t}\mathbf{1}\left(FS_{k-1}<FS_k\right)\left[\sum_{g\in\mathcal{G}_k^+}\mathbb{P}(g)-\sum_{g\in\mathcal{G}_k^-}\mathbb{P}(g)\right]\\
&\leq \sum_{k=2}^{t}\mathbf{1}\left(FS_{k-1}<FS_k\right)\sum_{g\in\mathcal{G}_k^+}\mathbb{P}(g) \qquad (25)\\
&\leq \sum_{k=2}^{t}\sum_{g\in\mathcal{G}_k^+}\mathbb{P}(g),
\end{aligned}$$

the condition is verified. Once more, the argument for the upper bound is analogous.

B Alternative Decomposition

Here we consider a decomposition for the per-period reduced form estimands using causal effects that compare treated potential outcomes of different treatment lengths. To do so, we extend our notation and define

$$\Delta_t^{\tau,\tau'}(g) \coloneqq \mathbb{E}\left[Y_{it}(1,\tau) - Y_{it}(1,\tau')|g\right],$$

where once more g specifies a history of IV latent types. For effects comparing the treated potential outcome of treatment length τ to the untreated potential outcome, we use the original notation with τ as the only superscript. Proposition 3.1′ below formalizes an alternative decomposition for the reduced forms based on this generalization of causal effects. This result is an adaptation of Proposition 3.1.

Proposition 3.1′. *Under Assumptions 3.1 and 3.2, for each $t \in \mathcal{T} \setminus \{1\}$,*

$$\begin{aligned}
RF_t &= \mathbb{P}\left(C_{1:t}\right)\Delta_t^{t-1}(C_{1:t}) \\
&\quad + \sum_{k=2}^{t} \mathbb{P}\left(C_{1:k-1}, AT_k\right)\Delta_t^{t-1,t-k}\left(C_{1:k-1}, AT_k\right) \\
&\quad + \sum_{k=2}^{t}\Bigg[\mathbb{P}\left(NT_{1:k-1}, C_{k:t}\right)\Delta_t^{t-k}\left(NT_{1:k-1}, C_{k:t}\right) \\
&\qquad\qquad + \sum_{\ell=k+1}^{t} \mathbb{P}\left(NT_{1:k-1}, C_{k:\ell-1}, AT_\ell\right)\Delta_t^{t-k,t-\ell}\left(NT_{1:k-1}, C_{k:\ell-1}, AT_\ell\right)\Bigg] \\
&\quad - \sum_{k=2}^{t}\Bigg[\mathbb{P}\left(NT_{1:k-1}, F_{k:t}\right)\Delta_t^{t-k}\left(NT_{1:k-1}, F_{k:t}\right) \\
&\qquad\qquad + \sum_{\ell=k+1}^{t} \mathbb{P}\left(NT_{1:k-1}, F_{k:\ell-1}, AT_\ell\right)\Delta_t^{t-k,t-\ell}\left(NT_{1:k-1}, F_{k:\ell-1}, AT_\ell\right)\Bigg]
\end{aligned} \tag{26}$$

under the convention that $\sum_{\ell=t+1}^{t} \dots = 0$.

Proof. Working the products in Equation 19 we get:

$$\begin{aligned}
\mathbb{E}\left[Y_{i,t}|Z_i = 0, C_1\right] &= \mathbb{P}\left(C_{1:t}|C_1\right)\mathbb{E}\left[Y_{i,t}(0)|C_{1:t}\right] \\
&\quad + \sum_{k=2}^{t} \mathbb{P}\left(C_{1:k-1}, AT_k|C_1\right)\mathbb{E}\left[Y_{i,t}(1, t-k)|C_{1:k-1}, AT_k\right].
\end{aligned}$$

The analogous expression for $\mathbb{E}\left[Y_{i,t}|Z_i = 1, C_1\right]$ follows from decomposing $\mathbb{E}\left[Y_{i,t}(1, t-1)|C_1\right]$

(Equation 18) in all the IV latent histories a C_1 can take:

$$\mathbb{E}\left[Y_{i,t}|Z_i=1,C_1\right] = \mathbb{P}\left(C_{1:t}|C_1\right)\mathbb{E}\left[Y_{i,t}(1,t-1)|C_{1:t}\right] + \sum_{k=2}^{t}\mathbb{P}\left(C_{1:k-1},AT_k|C_1\right)\mathbb{E}\left[Y_{i,t}(1,t-1)|C_{1:k-1},AT_k\right].$$

Thus,

$$\mathbb{E}\left[Y_{i,t}|Z_i=1,C_1\right] - \mathbb{E}\left[Y_{i,t}|Z_i=0,C_1\right] = \mathbb{P}\left(C_{1:t}|C_1\right)\Delta_t^{t-1}(C_{1:t}) + \sum_{k=2}^{t}\mathbb{P}\left(C_{1:k-1},AT_k|C_1\right)\Delta_t^{t-1,t-k}\left(C_{1:k-1},AT_k\right).$$

As in the proof of Proposition 3.1 (see Appendix A.1), the desired result follows from noting that for all other relevant latent groups, the dynamic of treatment access is the same as for C1 units. Also, it is necessary to account for the probability of each history. □

With this decomposition, if we assume no defiance in all periods, it is possible to interpret the reduced form and IV estimands as a positively weighted average of causal effects. The difference is on the interpretation of the causal effects considered. This decomposition shows, for example, that it is possible to have a negative reduced form estimand even when all individuals are better off (compared to the untreated scenario) when treatment effects fade out with time since treatment. To see this, consider RF_2 assuming $\mathbb{P}(NT_1,F_2)=0$:

$$RF_2 = \mathbb{P}(C_1,C_2)\Delta_2^1(C_1,C_2) + \mathbb{P}(C_1,AT_2)\Delta_2^{1,0}(C_1,AT_2) + \mathbb{P}(NT_1,C_2)\Delta_2^0(NT_1,C_2),$$

which may be negative only because $\Delta_2^{1,0}(C_1,AT_2)<0$, even when the causal effects comparing treated and untreated potential outcomes are positive for all latent groups. For this reason, we believe that an interpretation based on Equation 26 is challenging even though the decomposition suggests that all causal effects are positively weighted.

Additionally, if we consider the leading effect in Equation 26 as the effect for $C_{1:t}$ units, an interpretation based on this decomposition is further challenged. After all, the latent group of the leading effect changes with time and is for a smaller population as time passes. On the other hand, in the decomposition of Proposition 3.1, we have that the leading effect is for the same latent group in all periods, the C_1 units.

C Partial identification with general bounds on treatment effects

In this section, we state a version of Proposition 3.3 in which we do not require the lower bound for the treatment effects to be nonpositive and the upper bound to be nonnegative. Proposition 3.3′ only considers the more general version of the bounds (Equations 14 and 15) because the bounds in Equations 16 and 17 are generally valid (and continue to be weakly tighter). The proof of this proposition is provided in Appendix A.4.

Proposition 3.3′. *Suppose Assumptions 3.1 and 3.2 hold. If, for $t \in \mathcal{T} \setminus \{1\}$, there exist $\underline{\Delta}_t, \overline{\Delta}_t \in \mathbb{R}$ such that, for all $\tau \in \{0, ..., t-2\}$, if $g \in \mathcal{G}_{t-\tau}$ and $\mathbb{P}(g) > 0$, $\underline{\Delta}_t \leq \Delta_t^{\tau}(g) \leq \overline{\Delta}_t$, then a lower bound for $\Delta_t^{t-1}(C_1)$ is given by*

$$\begin{aligned}\frac{RF_t}{FS_1} &+ \mathbf{1}\left(\underline{\Delta}_t < 0\right)\mathbb{P}\left(D_{i,t} > D_{i,1}|Z_i = 0\right)\frac{\underline{\Delta}_t}{FS_1} + \mathbf{1}\left(\underline{\Delta}_t \geq 0\right)\max\left\{FS_1 - FS_t, 0\right\}\frac{\underline{\Delta}_t}{FS_1} \\ &- \mathbf{1}\left(\overline{\Delta}_t \geq 0\right)\mathbb{P}\left(D_{i,t} > D_{i,1}|Z_i = 1\right)\frac{\overline{\Delta}_t}{FS_1} - \mathbf{1}\left(\overline{\Delta}_t < 0\right)\max\left\{FS_t - FS_1, 0\right\}\frac{\overline{\Delta}_t}{FS_1}\end{aligned}$$

and an upper bound is given by

$$\begin{aligned}\frac{RF_t}{FS_1} &+ \mathbf{1}\left(\overline{\Delta}_t \geq 0\right)\mathbb{P}\left(D_{i,t} > D_{i,1}|Z_i = 0\right)\frac{\overline{\Delta}_t}{FS_1} + \mathbf{1}\left(\overline{\Delta}_t < 0\right)\max\left\{FS_1 - FS_t, 0\right\}\frac{\overline{\Delta}_t}{FS_1} \\ &- \mathbf{1}\left(\underline{\Delta}_t < 0\right)\mathbb{P}\left(D_{i,t} > D_{i,1}|Z_i = 1\right)\frac{\underline{\Delta}_t}{FS_1} - \mathbf{1}\left(\underline{\Delta}_t \geq 0\right)\max\left\{FS_t - FS_1, 0\right\}\frac{\underline{\Delta}_t}{FS_1}.\end{aligned}$$