Longitudinal weighted and trimmed treatment effects with flip interventions

Alec McClean[1*], Iván Díaz[1], and Alex Levis[2]

[1]Division of Biostatistics, New York University Grossman School of Medicine
[2]Department of Statistics and Data Science, Carnegie Mellon University

Preliminary draft: comments welcome

Abstract

Weighting and trimming are popular methods for addressing positivity violations in causal inference. While well-studied with single-timepoint data, standard methods do not easily generalize to non-baseline covariates in longitudinal data. Consequently, these effects remain susceptible to positivity violations at subsequent timepoints. In this paper, we extend weighting and trimming to longitudinal data via dynamic stochastic interventions. We introduce "flip" interventions, which maintain the treatment status of subjects who would have received the target treatment, and flip others' treatment to the target with probability equal to their weight (e.g., overlap weight, trimming indicator). With single-timepoint data, we show that a large class of weighted average treatment effects are equivalent to "interventional" flip effects, which are the difference in mean potential outcomes under a pair of flip interventions standardized by the mean difference in number of treated units. Thus, in single-timepoint data, weighted effects can be ascribed a novel policy interpretation via flip interventions. With longitudinal data, flip interventions provide interpretable weighting or trimming on non-baseline covariates. Crucially, flip interventions are policy-relevant since they could be implemented in practice. By contrast, we show that other approaches for weighting or trimming on non-baseline covariates do not retain this property. We derive efficient estimators based on efficient influence functions when the weight is a smooth function of the propensity score. We construct multiply robust-style and sequentially doubly robust-style estimators that achieve root-n consistency and asymptotic normality under nonparametric conditions.

Keywords: *Causal inference; longitudinal data; positivity violations; weighting; trimming; dynamic stochastic interventions; nonparametrics*

*Corresponding author: `hadera01@nyu.edu`

1 Introduction

There is a large and growing literature in causal inference for estimating treatment effects from observational data under violations of the positivity assumption. For binary treatments, positivity requires subjects in every covariate stratum to have non-zero probability of receiving both treatment and control [Hernán and Robins, 2020]. In longitudinal settings, positivity requires non-zero probability for every treatment regime under consideration, which can become increasingly difficult to satisfy with more timepoints as the number of treatment regimes increases exponentially. Full positivity violations (zero probability of certain regimes) render causal effects unidentifiable, while "practical" violations (near-zero probability of certain regimes) inflate variance estimates. Both hamper meaningful scientific conclusions from data [Kang and Schafer, 2007, Moore et al., 2012, Petersen et al., 2012].

Several approaches have been developed to define effects robust to positivity violations. We focus on weighting/trimming and dynamic stochastic interventions. In single-timepoint data, weighted average treatment effects (WATEs) take the form

$$\mathbb{E}\left[\frac{\mathbb{E}\{Y(1)-Y(0)\mid X\}f(X)}{\mathbb{E}\{f(X)\}}\right]$$

where $Y(1)$ and $Y(0)$ are potential outcomes, X are covariates, and $f(X)$ is a weight, which is typically based on the propensity score $\pi(X) := \mathbb{P}(A = 1 \mid X)$. Examples include overlap weights [Li et al., 2018], entropy weights [Hainmueller, 2012], and trimmed or smooth trimmed effects [Crump et al., 2009, Yang and Ding, 2018]. WATEs can also be defined implicitly post hoc by estimating weights $f(X)$ to directly maximize covariate balance between treatment and control groups and then estimating the identified analog of the WATE in the display above [Cohn et al., 2023].

Weighting and trimming are well-studied for single-timepoint data, with extensions beyond binary treatment (e.g., Branson et al. [2023], Li and Li [2019]), but their extension to longitudinal settings remains challenging. Current approaches typically trim using only baseline covariates [Jensen et al., 2024, Petersen et al., 2012], leaving them vulnerable to positivity violations from non-baseline covariates. While some longitudinal weighting methods exist (e.g, Zeng et al. [2023]), the counterfactual estimands they target remains unclear.

Meanwhile, dynamic stochastic interventions offer an alternative approach by shifting treatment probabilities and adapting to positivity violations at each timepoint. Pioneered by Robins et al. [2004] and Stock [1989], these interventions have evolved through various formulations including modified treatment policies (MTPs) [Díaz and van der Laan, 2012, Haneuse and Rotnitzky, 2013], threshold interventions [Taubman et al., 2009], incremental propensity score interventions [Kennedy, 2019], and recent extensions to complex settings [Díaz et al., 2023, McClean et al., 2024b, Schindl et al., 2024, Stensrud et al., 2024]. The natural applicability of dynamic stochastic interventions to longitudinal data has made them increasingly popular.

1.1 Structure of the paper and our contributions

Methods for weighting/trimming and dynamic stochastic interventions have developed largely in parallel. In Section 2, we demonstrate that WATEs correspond to differences in potential outcomes under pairs of "flip" interventions. Each pair contains one intervention targeting treatment and one targeting control. If a subject would have taken the target treatment, the flip intervention does not intervene; otherwise, it flips the subject to the target treatment with probability equal to the weight $f(X)$. We show that WATEs are *interventional* effects. They are the average effect of the flip intervention targeting treatment compared to the flip intervention targeting control, per unit of additional treatment [Zhou and Opacic, 2022]. This connection may be of independent interest because it yields a direct policy interpretation for a large class of WATEs, including those that balance covariates directly and define the WATE post hoc.

Building on this insight, Section 3 introduces our notation for longitudinal data, while Section 4 extends flip interventions to longitudinal data. These interventions flip subjects toward a target treatment regimen using weights/flipping probabilities constructed from non-baseline covariates. Crucially, they remain identifiable even under positivity violations at all timepoints and are *single-world* interventions, relying only on data observable under the intervention itself, and therefore are policy-relevant because they could be implemented in practice [Richardson and Robins, 2013]. We define *longitudinal* interventional flip effects, which capture the mean difference in potential outcomes under two flip interventions, standardized by the average change in the number of treatments per timepoint. In Section 4.2, we investigate the properties of longitudinal interventional flip effects and outline issues with alternative estimands one might consider. Crucially, we show that simple extensions of weighting and trimming to longitudinal data do not correspond to interventions that could be implemented in practice, justifying our focus on flip interventions.

Section 5 develops efficient estimators for flip effects when the flipping probability is a smooth function of the propensity score. We present multiply robust and sequentially doubly robust estimators based on efficient influence functions that achieve parametric convergence rates under nonparametric conditions. We establish two new results: (1) tighter bounds on the bias of multiply robust estimators using the minimum of two error decompositions, and (2) the first sequentially doubly robust guarantee for a dynamic stochastic intervention that depends on unknown weights. For single-timepoint data, this result may be of independent interest because it provides a doubly robust-style estimator for a large class of WATEs. In future work, we will illustrate our methods in `R` with simulations and data analyses.

1.2 Mathematical notation

For a function $f(Z)$, we use $\|f\| = \sqrt{\int f(z)^2 d\mathbb{P}(z)}$ to denote the $L_2(\mathbb{P})$ norm, $\mathbb{P}(f) = \int_{\mathcal{Z}} f(z)d\mathbb{P}(z)$ to denote the average with respect to the underlying distribution $\mathbb{P}$, and $\mathbb{P}_n(f) = \frac{1}{n}\sum_{i=1}^{n} f(Z_i)$ to denote the empirical average with respect to n observations. In a standard abuse of notation, when A is an event we let $\mathbb{P}(A)$ denote the probability of A. We also denote expectation and variance with respect to the underlying distribution by $\mathbb{E}$ and $\mathbb{V}$, respectively. We use $a \wedge b$ for minimum and $a \vee b$ for maximum, and $a \lesssim b$ to mean $a \leq Cb$ for some constant C. We use $\rightsquigarrow$ to denote convergence in distribution, and $\xrightarrow{p}$ for convergence in probability. Additionally, we use $o_{\mathbb{P}}(\cdot)$ to denote convergence in probability to zero, i.e., if X_n is a sequence of random variables then $X_n = o_{\mathbb{P}}(r_n)$ implies $\left|\frac{X_n}{r_n}\right| \xrightarrow{p} 0$.

2 Single-timepoint flip interventions and weighted average treatment effects

We assume data $\{(X_i, A_i, Y_i)\}_{i=1}^{n} \overset{iid}{\sim} \mathbb{P}$ where $X \in \mathbb{R}^d$ are covariates, $A \in \{0, 1\}$ is a binary treatment, and $Y \in \mathbb{R}$ is an outcome. Moreover, we assume that the observed data corresponds to complete data $\{(X_i, A_i, Y_i(0), Y_i(1)\}_{i=1}^{n} \overset{iid}{\sim} \mathbb{P}^c$ where $Y(a)$ is the potential outcome under treatment a. We let $Y(D)$ denote the potential outcome under treatment decision D, where D is a random variable that can depend on the treatment A and covariates X. Finally, we let $\pi(X) = \mathbb{P}(A = 1 \mid X)$ denote the propensity score.

When positivity is violated, a common estimand of interest is the weighted average treatment effect (WATE), which we denote as

$$\mathbb{E}\left[\frac{\mathbb{E}\{Y(1) - Y(0) \mid X\}f(X)}{\mathbb{E}\{f(X)\}}\right]. \tag{1}$$

Our main result in this section shows that a large class of WATEs can be defined via flip interventions. First, we define a pair of flip interventions using the weight function $f(X)$. Then, we define an interventional flip effect based on this pair. And finally, we establish that this interventional flip effect is exactly the WATE in (1).

Definition 1 (Single-timepoint flip interventions)**.** Suppose access to a weight function $f(X) : \mathbb{R}^d \to [0, 1]$, which maps from the covariate information to $[0, 1]$. Define a pair of flip interventions, one for each $a \in \{0, 1\}$, as

$$D_f(a) = \mathbb{1}(A = a)A + \mathbb{1}(A \neq a)\Big[a\mathbb{1}\{V \leq f(X)\} + A\mathbb{1}\{V > f(X)\}\Big]$$

where $V \sim \text{Unif}(0, 1)$. In words:

- if the treatment, A, equals the target treatment a, the flip intervention does nothing;
- otherwise, it flips the subject to the target treatment with probability $f(X)$.

Remark 1. In Definition 1, we introduce an auxiliary random variable V. This is a standard device in the literature on stochastic interventions, and captures the idea of randomly reassigning treatment according to a Bernoulli distribution with a modified probability [Díaz et al., 2023].

Next, we define an interventional flip effect based on a pair of flip interventions.

Definition 2 (Interventional flip effect)**.** For a pair of flip interventions $\{D_f(0), D_f(1)\}$ from Definition 1, we define the *interventional* flip effect as

$$\psi_f = \frac{\mathbb{E}\big[Y\{D_f(1)\} - Y\{D_f(0)\}\big]}{\mathbb{E}\{D_f(1) - D_f(0)\}}. \tag{2}$$

This is the average effect on potential outcomes of $D_f(1)$ compared to $D_f(0)$ *per unit of additional treatment*.

We define ψ_f as an *interventional* effect, which standardizes the mean difference in potential outcomes by the mean difference in the number of treated units. Therefore, ψ_f captures the notion of the treatment effect *per unit treated*. This type of effect is useful for understanding the effectiveness of one policy versus another while accounting for how many subjects are affected by the change in policy. It has been studied previously in the literature; see, e.g., Zhou and Opacic [2022] and references therein. Finally, we establish the interventional flip effect in (2) is exactly the WATE in (1).

Proposition 1 (Interventional flip effects are WATEs)**.** *Let ψ_f denote an interventional flip effect from Definition 2. Suppose $Y(a) \perp\!\!\!\perp A \mid X$ for $a \in \{0,1\}$. Then,*

$$\psi_f = \mathbb{E}\left[\frac{\mathbb{E}\{Y(1) - Y(0) \mid X\} f(X)}{\mathbb{E}\{f(X)\}}\right].$$

Proposition 1 establishes that interventional flip effects are equal to weighted average treatment effects. While we leverage this result to extend weighting and trimming to longitudinal data, this equivalence is itself be of independent interest. Crucially, Proposition 1 demonstrates that any WATE with weights bounded between zero and one corresponds to a contrast in specific interventions—a connection that, to our knowledge, has not been established previously. Table 1 illustrates several important examples, including the average treatment effect on the treated (ATT) or control (ATC), the average treatment effect on the overlap population (ATO), and the trimmed average treatment effect.[1] Furthermore,

[1]Table 1 is adapted from the helpful tutorial on weighting, here: `https://www2.stat.duke.edu/~fl35/OW/ICHPS2023.pdf`

this framework also allows researchers to retrospectively define the implicit estimand underlying any weighting estimator that directly balances covariates and constrains weights to lie within the interval $[0,1]$. For instance, if weights are constructed from a separate sample and subsequently applied to estimate a WATE in a new sample conditional on these estimated weights, the resulting estimand is equivalent to the interventional flip effect $\psi_{\widehat{f}}$, where the estimated weighting function $\widehat{f}$ is treated as fixed.

Remark 2. The flip intervention in Definition 1 depends on the observed treatment. As a result, Proposition 1 requires exchangeability, despite not being an identification result. This requirement can be relaxed by defining alternative flip interventions that do not depend on the observed treatment:

$$D_f(a) = \mathbb{1}\left\{V \leq \mathbb{1}(a=1)f(X) + \{1-f(X)\}\pi(X)\right\}.$$

With such an intervention, Proposition 1 would hold without exchangeability. However, one loses the interpretation of $D_f(a)$ as a "flip" intervention.

Estimand; ψ_f	**Weight/flipping probability;** $f(X)$
ATE	1
ATT	$\pi(X)$
ATC	$1-\pi(X)$
ATO	$\pi(X)(1-\pi(X))$
Trimmed ATE	$\mathbb{1}\{\varepsilon \leq \pi(X) \leq 1-\varepsilon\}$ for $\varepsilon \geq 0$
Smooth trimmed ATE	$S\{\pi(X);\varepsilon\}$ where $S(x;\varepsilon)$ approximates $\mathbb{1}(\varepsilon \leq x \leq 1-\varepsilon)$
Matching-weighted ATE	$\pi(X) \wedge 1-\pi(X)$
Direct covariate balancing	Varies and is data-dependent; derived to directly balance covariates

Table 1: Common weighted average treatment effect estimands and weights/flipping probabilities $f(X)$.

3 Setup and background for longitudinal flip interventions

As in the single-timepoint case, we assume n observations drawn iid from some distribution $\mathbb{P}$ in a space of distributions $\mathcal{P}$; i.e., we observe data $\{Z_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P} \in \mathcal{P}$. We assume each observation consists of longitudinal data over T timepoints, so that

$$Z = (X_1, A_1, X_2, A_2, \ldots, X_T, A_T, Y),$$

where $X_t \in \mathbb{R}^d$ are time-varying covariates (X_1 are baseline covariates), $A_t \in \{0,1\}$ is a time-varying binary treatment, and $Y \in \mathbb{R}$ is the ultimate outcome of interest. For a time-varying random variable O_t, let $\overline{O}_t = (O_1, \ldots, O_t)$ denote its history up to time t and $\underline{O}_t = (O_t, \ldots, O_T)$ denote its future from time t. Let $H_t = (\overline{X}_t, \overline{A}_{t-1})$ denote covariate and treatment history up until treatment in timepoint t.

We formalize the definition of causal effects using a nonparametric structural equation model (NPSEM) [Pearl, 2009]. We assume the existence of deterministic functions $\{f_{X,t}, f_{A,t}\}_{t=1}^{T}$ and f_Y such that

$$\begin{aligned} X_t &= f_{X,t}(A_{t-1}, H_{t-1}, U_{X,t}), \\ A_t &= f_{A,t}(H_t, U_{A,t}), \text{ and} \\ Y &= f_Y(A_T, H_T, U_Y). \end{aligned}$$

Here, $\Big\{\big\{U_{X,t}, U_{A,t} : t \in \{1, \ldots, T\}\big\}, U_Y\Big\}$ is a vector of exogenous variables. Subsequently, we'll define restrictions on their joint distribution that facilitate identification of causal effects. We will define the effects in terms of hypothetical interventions in which equation $A_t = f_{A,t}(H_t, U_{A,t})$ is removed from the structural model and the exposure is assigned as a new random variable D_t (which could be deterministic). An intervention that sets exposures up to time $t-1$ to $\overline{D}_{t-1} \equiv \{D_1, \ldots, D_{t-1}\}$ generates counterfactual variables $X_t(\overline{D}_{t-1}) = f_{X,t}\big\{D_{t-1}, H_{t-1}(\overline{D}_{t-2}), U_{X,t}\big\}$ and $A_t(\overline{D}_{t-1}) = f_{A,t}\big\{H_t(\overline{D}_{t-1}), U_{A,t}\big\}$, where the counterfactual history is defined recursively as $H_t(\overline{D}_{t-1}) = \big\{\overline{D}_{t-1}, \overline{X}_t(\overline{D}_{t-1})\big\}$ and $A_1(D_0) = A_1$ and $X_1(D_0) = X_1$. The variable $A_t(\overline{D}_{t-1})$ is called the *natural value of treatment* [Richardson and Robins, 2013, Young et al., 2014], and represents the possibly counterfactual value of treatment that would have been observed at time t under an intervention carried out up to time $t-1$ but discontinued thereafter. An intervention in which all treatment variables up to $t = T$ are intervened on generates a counterfactual outcome $Y\left(\overline{D}_T\right) = f_Y\big\{D_T, H_T(\overline{D}_{T-1}), U_Y\big\}$.

3.1 Causal assumptions

The NPSEM implicitly contains the consistency assumption because one subject's information does not depend on another's. This assumption would be violated if there were interference between subjects [Tchetgen Tchetgen and VanderWeele, 2012]. We consider two exchangeability assumptions on the exogeneous variables.

Assumption 1 (Standard sequential randomization). $U_{A,t} \perp\!\!\!\perp \big\{\underline{U}_{X,t+1}, U_Y\big\} \mid H_t$ for all $t \leq T$.

Assumption 2 (Strong sequential randomization). $U_{A,t} \perp\!\!\!\perp \big\{\underline{U}_{X,t+1}, \underline{U}_{A,t+1}, U_Y\big\} \mid H_t$ for all $t \leq T$.

Assumption 1 is standard for the identification of effects under dynamic stochastic interventions [Díaz et al., 2023]. It is satisfied if the common causes of the treatment A_t and future covariates are measured. Assumption 2 is stronger. It is satisfied if common causes of treatment A_t and future covariates *and treatments* are measured. This assumption is similar to that required by Richardson and Robins [2013] (cf. Theorem 31), and allows identification of effects under certain interventions that depend on the natural value of treatment [Young et al., 2014]. Finally, note that we do not immediately require the typical positivity assumption, which says that time-varying propensity scores are bounded away from zero and one ($0 < \mathbb{P}(A_t = 1 \mid H_t) < 1$), because we will construct interventions which adapt to positivity violations. Some flip interventions will require a version of positivity, and we will introduce the assumption when it is needed.

4 Longitudinal weighting and trimming with flip interventions

In this section, we extend flip interventions and interventional flip effects from Section 2 to longitudinal data. We define longitudinal flip interventions that perform weighting or trimming on non-baseline covariates, and then develop longitudinal analogs to interventional flip effects. Depending on the weighting function, these effects remain identifiable even with arbitrary positivity violations. Importantly, the flip interventions we consider are "single-world", relying only on information observable under the series of interventions, which preserves their policy relevance for practical implementation.

To conclude the section, we investigate properties of longitudinal intervention flip effects and outline limitations of alternative estimands. Due to their single-world nature, longitudinal interventional flip effects can be driven both by differences in potential outcomes under different regimes and by the impact of earlier interventions on subsequent covariates and natural treatment values. While one might prefer estimands that isolate mechanistic differences in potential outcomes, we demonstrate that such estimands would require cross-world and future information to guide earlier interventions, creating significant practical and interpretational challenges.

4.1 Single-world flip interventions

We now propose flip interventions that extend the interventions in Section 2 to longitudinal data. We begin by defining flip interventions targeting a specific longitudinal regime and then longitudinal interventional effects contrasting two flip interventions. We establish conditions under which the resulting effects remain identifiable without standard positivity assumptions.

Definition 3 (Single-world flip interventions)**.** Suppose access to some function f_t^c : $\{0,1\} \times \mathcal{H}_t \to [0,1]$, which maps a target treatment and covariate history to $[0,1]$. Let $\overline{a}_T = \{a_1, \ldots, a_T\} \in \{0,1\}^T$ be the target regime. A *flip intervention* at time t targeting

a_t is

$$D_t = \begin{cases} A_t(\overline{D}_{t-1}), & \text{if } A_t(\overline{D}_{t-1}) = a_t, \\ a_t \mathbb{1}\Big[V_t \leq f_t^c\{a_t; H_t(\overline{D}_{t-1})\}\Big] + A_t(\overline{D}_{t-1})\mathbb{1}\big[V_t > f_t^c\{a_t; H_t(\overline{D}_{t-1})\}\big], & \text{otherwise,} \end{cases}$$

where $V_1, \ldots, V_T$ are iid Unif$(0, 1)$ random variables with $\{V_1, \ldots, V_T\} \perp\!\!\!\perp Z$.

In words, at time t:

- if the natural value of treatment is already a_t, the flip intervention does nothing;
- otherwise, it flips the subject to the target treatment with probability $f_t^c\{a_t; H_t(\overline{D}_{t-1})\}$.

Importantly, the intervention is "single-world," meaning the decision at time t depends only on information observable under interventions performed up to that point ($\overline{D}_{t-1}$). A related nuance is that the weighting function can be counterfactual because it can depend on the natural value of treatment. For example, if $f_t^c\{a_t; H_t(\overline{D}_{t-1})\}$ is a trimming indicator, it will be a trimming indicator using the "natural" propensity score, i.e., $f_t^c\{a_t; H_t(\overline{D}_{t-1})\} = \mathbb{1}\big[\varepsilon < \mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\} < 1 - \varepsilon\big]$. Therefore, both the function and its argument are counterfactual, but would be observed under the series of interventions $\overline{D}_{t-1}$. The superscript "c" emphasizes the "counterfactual" nature of $f_t^c(\cdot)$. Table 2 gives additional examples. Next, we define a longitudinal interventional flip effect.

Definition 4 (Longitudinal interventional flip effect)**.** Let $\overline{D}_T$ and $\overline{D}'_T$ denote two flip interventions targeting $\overline{a}_T$ and $\overline{a}'_T$, respectively. Define the longitudinal interventional flip effect from these two interventions as

$$\frac{\mathbb{E}\Big\{Y(\overline{D}_T) - Y(\overline{D}'_T)\Big\}}{\frac{1}{T}\sum_{t=1}^T |\mathbb{E}\left(D_t - D'_t\right)|}. \tag{3}$$

The effect defined in (3) generalizes the single-timepoint interventional flip effect, characterizing the *average change in potential outcomes standardized by the average absolute per-timepoint change in the number of treatments.* While the numerator matches the single-timepoint definition, multiple options exist for the denominator. We adopt a per-timepoint absolute difference approach, allowing for flexible policy contrasts where $\mathbb{E}(D_t - D'_t)$ may change signs across timepoints. In scenarios with monotonic interventions—where, without loss of generality, $D_t \geq D'_t$ almost surely for all $t \in \{1, \ldots, T\}$—the absolute value function becomes unnecessary, and one retains the same interpretation as in the single-timepoint case.

Remark 3. There are other options for defining the denominator, each corresponding to different measures of distance between two longitudinal treatment distributions. We leave a full investigation to future work, but briefly note that other options include:

Type of weighting	Weight/flipping probability $f_t^c\{a_t; H_t(\overline{D}_{t-1})\}$
No weighting	1
Weighting to subjects that can take target treatment	$\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}$
Weighting to subjects that can take non-target treatment	$1 - \mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}$
Overlap weighting	$\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}\Big[1 - \mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}\Big]$
Trimming	$\mathbb{1}\{\varepsilon \leq \mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\} \leq 1 - \varepsilon\}$ for $\varepsilon \geq 0$
Smooth trimming	$S\big[\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}; \varepsilon\big]$, where $S(x; \varepsilon)$ approximates $\mathbb{1}(\varepsilon \leq x \leq 1 - \varepsilon)$
Matching-style weighting	$\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\} \wedge \big(1 - \mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}\big)$
Direct covariate balancing	Varies and is data-dependent; these must be derived to directly balance *counterfactual* covariates

Table 2: Extending weights/flipping probabilities to longitudinal data.

- *Average per-timepoint probability of switching treatment assignment*: $\frac{1}{T}\sum_{t=1}^{T} \mathbb{P}(D_t \neq D_t')$. This is an intuitive notion of the distance between the two treatment distributions, but one should note that it would not yield interventional flip effects in single-timepoint data. Nonetheless, it is a useful alternative to our proposal in (3). Moreover, with binary treatment, note also that $\mathbb{P}(D_t \neq D_t') = \mathbb{E}(D_t)\{1 - \mathbb{E}(D_t')\} + \{1 - \mathbb{E}(D_t)\}\mathbb{E}(D_t')$. Therefore, one could use this switching distance instead of the denominator in (3) and our subsequent results for efficient estimation would apply.

- *Joint distributional distances*: To incorporate dependence across timepoints explicitly, one might use distances between joint distributions, such as f-divergences or optimal transport metrics, comparing the distributions $\mathcal{P}$ and $\mathcal{P}'$ of $D_1, \ldots, D_T$ and $D_1', \ldots, D_T'$, respectively. Depending on the choice of distance, estimation may be more complex.

The next result establishes conditions under which flip interventions correspond to identifiable functionals, even under arbitrary positivity violations. We focus on the mean potential outcome under one intervention, $\mathbb{E}\{Y(\overline{D}_T)\}$, and the average treatment at an arbitrary timepoint, $\mathbb{E}(D_t)$.

Theorem 1 (Longitudinal identification with flip interventions). *Let $\overline{D}_T = \{D_1, D_2, \ldots, D_T\}$ denote flip interventions as in Definition 3 targeting treatment regime $\overline{a}_T$. Moreover, let $f_t(a_t; H_t)$ denote the identified analog to f_t^c, which replaces $\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}$ by $\mathbb{P}(A_t = a_t \mid H_t)$ in Table 2. Then, suppose the NPSEM and Assumption 2 hold and f_t^c and f_t are constructed such that*

$$\mathbb{P}(A_t = a_t \mid H_t) = 0 \implies f_t(a_t; H_t) = 0$$

or positivity holds such that $\mathbb{P}\big\{\mathbb{P}(A_t = a_t \mid H_t) = 0\big\} = 0$. Then,

$$\begin{aligned}
\mathbb{E}\big\{Y(\overline{D}_T)\big\} &= \sum_{\overline{b}_T \in \{0,1\}^T} \int_{\overline{\mathcal{X}}_T} \mathbb{E}\big(Y \mid \overline{A}_T = \overline{b}_T, \overline{X}_T = \overline{x}_T\big) \prod_{t=1}^{T} Q_t\big(b_t \mid \overline{b}_{t-1}, \overline{x}_t\big) d\mathbb{P}(x_t \mid \overline{b}_{t-1}, \overline{x}_{t-1}) \\
&= \mathbb{E}\left[Y \prod_{t=1}^{T} \frac{Q_t(A_t \mid H_t)}{\mathbb{P}(A_t \mid H_t)}\right] \qquad (4)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\big(D_t\big) &= \sum_{\overline{b}_{t-1} \in \{0,1\}^{t-1}} \int_{\overline{\mathcal{X}}_t} Q_t(1 \mid \overline{b}_{t-1}, \overline{x}_t) d\mathbb{P}(x_t \mid \overline{b}_{t-1}, \overline{x}_{t-1}) \cdot \prod_{s=1}^{t-1} Q_s(b_s \mid \overline{b}_{s-1}, \overline{x}_s) d\mathbb{P}(x_s \mid \overline{b}_{s-1}, \overline{x}_{s-1}) \\
&= \mathbb{E}\left[Q_t(A_t = 1 \mid H_t) \prod_{s=1}^{t-1} \frac{Q_s(A_s \mid H_s)}{\mathbb{P}(A_s \mid H_s)}\right] \qquad (5)
\end{aligned}$$

where the probability of receiving the target treatment at time t is

$$Q_t(a_t \mid h_t) = \mathbb{P}(A_t = a \mid h_t) + f_t(a_t; h_t)\{1 - \mathbb{P}(A_t = a \mid h_t)\}.$$

Theorem 1 establishes that mean potential outcomes under flip interventions are identifiable under only strong sequential randomization and consistency. We provide g-formula identification [Robins, 1986] and inverse weighting identification in (4). We also derive a result for the average number of treatments, in (5). The result requires that the identified analog of the weighting function is zero when the observed propensity score for the target treatment is zero. This can be enforced by construction. For example, the overlap weights, trimming and smooth trimming weights, and matching-style weights in Table 2 all satisfy this condition. Other weights, like Shannon's entropy weights, also satisfy this condition. However, "no weighting" or weighting towards subjects that can take the non-target treatment would fail to satisfy this condition, and then it is necessary that positivity is satisfied with respect to the target treatment.

Before proceeding, we highlight several key observations that provide further context to these effects:

1. **Robustness to positivity violations.** Flip effects retain robustness to arbitrary positivity violations, making them a complementary alternative to incremental propensity score interventions (IPSIs) [Bonvini et al., 2023, Kennedy, 2019]. While both IPSIs and flip interventions are time-varying dynamic stochastic interventions that remain identifiable under positivity violations, flip interventions are distinct in being explicitly constructed to target specific treatment regimes.

2. **Interventions depending on the natural value of treatment.** A critique of interventions that depend on the *natural* value of treatment is that they may be impractical because this value is unobserved in practice. This issue can be addressed in two ways:

   (i) An approximation can be constructed by defining interventions based on a subject's intended treatment, which may closely approximate their natural treatment value. See Young et al. [2014, Section 6] for a discussion.

   (ii) It is possible to define flip interventions that do not depend on the natural value of treatment while still yielding the same identification result as Theorem 1. The next point elaborates on this modification.

3. **Relaxing the sequential randomization assumption.** The identification result in Theorem 1 relies on strong sequential randomization (Assumption 2) because flip interventions depend on the natural treatment value to retain an intuitive "flipping" interpretation. However, this assumption can be relaxed to standard sequential randomization (Assumption 1) by redefining the interventions so they do not depend on the natural treatment value. Specifically, one could instead define

$$D_t = \mathbb{1}\Big(V_t \leq \mathbb{1}(a_t = 1) f_t^c\{a_t; H_t(\overline{D}_{t-1})\} + \Big[1 - f_t^c\{a_t; H_t(\overline{D}_{t-1})\}\Big] \mathbb{P}\{A_t(\overline{D}_{t-1}) = 1 \mid H_t(\overline{D}_{t-1})\}\Big)$$

   where $V_t \sim \text{Unif}(0,1)$. These redefined interventions satisfy the identification result in (4) under standard sequential randomization and the identification result for the average number of treatments under only the consistency assumption embedded in the NPSEM. Moreover, they do not suffer from the practical concerns discussed in the previous point.

4. **Connections to maximally coupled policies.** Flip interventions that depend on the natural value of treatment are related to "maximally coupled generalized policies" [Levis et al., 2024], which minimize the number of subjects intervened on while preserving a target interventional propensity score, $Q_t(A_t \mid H_t)$. This approach was originally proposed to minimize bounds on causal effects under unmeasured confounding (e.g., adapting IPSIs [Levis et al., 2024, Section 3.3]). Here, we repurpose

these interventions because they have a nice interpretation as flip interventions. Examining their robustness to unmeasured confounding remains an open question for future work.

4.2 Properties of interventional flip effects; drawbacks of naive weighting or trimming

In this section, we investigate the properties of longitudinal interventional flip effects and outline issues with alternative estimands one might consider, justifying our focus on flip interventions. First, we note that longitudinal interventional flip effects satisfy a minimal property: if the treatment has no effect on the outcome but the interventions shift the treatment distribution, then the interventional flip effect is zero.

Proposition 2. *Let $\overline{D}_T$ and $\overline{D}'_T$ denote two flip interventions. If $Y(\overline{b}_T) = Y(\overline{b}'_T)$ for all $\overline{b}_T, \overline{b}'_T \in \{0,1\}^T$ but $\sum_{t=1}^T |\mathbb{E}(D_t) - \mathbb{E}(D'_t)| > 0$, then $\frac{\mathbb{E}\left\{Y(\overline{D}_T) - Y(\overline{D}'_T)\right\}}{T^{-1}\sum_{t=1}^T |\mathbb{E}(D_t - D'_t)|} = 0$.*

However, there is also a subtle consequence of the *single-world* nature of flip interventions: longitudinal interventional flip effects may reflect both mechanistic differences in potential outcomes and differences in sequential weighting/trimming. This occurs because the intervention $\overline{D}_t$ affects $X_{t+1}(\overline{D}_t)$ and $A_{t+1}(\overline{D}_t)$ as well as the ultimate outcome $Y(\overline{D}_T)$. As a result, only in certain extreme scenarios is it possible to isolate only the mechanistic difference $Y(\overline{a}_T) - Y(\overline{a}'_T)$ multiplied by a weight.

By consider, we can consider a general weighted effect that succeeds in isolating the weighted difference $Y(\overline{a}_T) - Y(\overline{a}'_T)$:

$$\mathbb{E}\left(\frac{\{Y(\overline{a}_T) - Y(\overline{a}'_T)\}\prod_{t=1}^T f_t^c\{a_t; H_t(\overline{a}_{t-1})\} f_t^c\{a'_t; H_t(\overline{a}'_{t-1})\}}{\mathbb{E}\left[\prod_{t=1}^T f_t^c\{a_t; H_t(\overline{a}_{t-1})\} f_t^c\{a'_t; H_t(\overline{a}'_{t-1})\}\right]}\right). \tag{6}$$

This is the weighted average treatment effect of $\overline{a}_T$ versus $\overline{a}'_T$, where the weights are based on the sequence of counterfactual propensity scores under each intervention over the whole regime. This effect isolates the difference $Y(\overline{a}_T) - Y(\overline{a}'_T)$ among subjects with non-zero weights for both regimes, and therefore may be considered a desirable estimand. However, it has two major limitations. First, it is "cross-world." Notice that the weighting function $f_t^c\{a_t; H_t(\overline{a}_{t-1})\} f_t^c\{a'_t; H_t(\overline{a}'_{t-1})\}$ depends on counterfactual covariates under two treatment regimes, and therefore one will be unobservable. Consequently, an intervention corresponding to this effect cannot be implemented as a single-world intervention and it cannot be falsified experimentally or implemented in practice. This limitation parallels natural effects in mediation [Andrews and Didelez, 2021, Richardson and Robins, 2013]. Second, the effect corresponds to a contrast under *future-dependent* interventions, as clarified by Proposition 3, next.

Proposition 3. *Let* $\Pi_T := \prod_{t=1}^{T} f_t^c\{a_t; H_t(\overline{a}_{t-1})\} f_t^c\{a_t'; H_t(\overline{a}_{t-1}')\}$ *and suppose* $a_t \geq a_t'$ *for all* $t \in \{1, \ldots, T\}$*. Then, the treatment decisions* $\overline{D}_T = \mathbb{1}(\overline{A}_T = \overline{a}_T)\overline{A}_T + \mathbb{1}(\overline{A}_T \neq \overline{a}_T)\Big\{\overline{a}_T \mathbb{1}\big(V \leq \Pi_T\big) + \overline{A}_T \mathbb{1}\big(V > \Pi_T\big)\Big\}$ *and* $\overline{D}_T' = \mathbb{1}(\overline{A}_T = \overline{a}_T')\overline{A}_T + \mathbb{1}(\overline{A}_T \neq \overline{a}_T')\Big\{\overline{a}_T' \mathbb{1}\big(V \leq \Pi_T\big) + \overline{A}_T\big(V > \Pi_T\big)\Big\}$ *yield a longitudinal interventional flip effect equal to the weighted treatment effect in* (6)*.*

Proposition 3 shows that the weighted effect in (6) is defined using *simultaneous* flip interventions: for subjects that would have followed the target regime, there is no intervention; otherwise, for those that would be in the trimmed set, the intervention flips their treatment to the target regime with probability equal to the weighted product across all timepoints. A key limitation of these effects arises from the nature of these interventions, which *simultaneously* alter the entire regime and rely on *future information* at earlier timepoints. For instance, the intervention at the first timepoint depends on a subject's natural treatment and covariate values at all timepoints. The simultaneous and cross-world nature of the intervention hampers the interpretability and practicality of the effects. These issues motivate our focus on flip interventions with longitudinal data and longitudinal interventional flip effects.

5 Estimation and inference

In this section, we outline methods for estimating flip effects. We focus on estimating mean potential outcomes $\mathbb{E}\big\{Y(\overline{D}_T)\big\}$ because estimating the average number of treatments $\mathbb{E}(D_t)$ follows the same approach. We develop methods for estimating $\mathbb{E}(D_t)$ in detail in Appendix A. Estimating and conducting inference on longitudinal interventional flip effects will follow by the delta under mild regularity conditions.

Throughout, we have assumed that the weight function was known a priori. We will continue to do so in this section. When this is not the case — for example, if one wanted to decide a trimming threshold or smooth trimming parameter data-adaptively — then estimation and inference are more complex; see Khan and Ugander [2022] for a review.

Moreover, we will assume the weight function is a smooth function of the propensity score. Specifically, we will assume the identified weight function $f_t(a_t; H_t)$ in Theorem 1 satisfies

$$f_t(a_t; H_t) = s_t\{\mathbb{P}(A_t = a_t \mid H_t)\}$$

where $s_t(\cdot)$ is twice differentiable with non-zero and bounded derivatives. This includes all the examples in Table 2 except trimming and matching-style weighting. The smoothness of the weight function is crucial to allow for $\sqrt{n}$-convergence under nonparametric conditions. With non-smooth weights, such as the trimming indicator, the lack of pathwise differentiability due to the non-smoothness of the weight function creates complications. Without additional assumptions, the performance of estimators for these effects is dictated by the behavior of propensity score estimators within the trimming indicator. While $\sqrt{n}$-rate estimation or valid inference may be possible under parametric models for the propensity scores or with specific nonparametric assumptions and estimators, general guarantees are unavailable. Therefore, we will focus on smooth weights, which are pathwise differentiable and allow for the construction of $\sqrt{n}$-consistent and asymptotically normal estimators under nonparametric assumptions by leveraging nonparametric efficiency theory and efficient influence functions [Bickel et al., 1993]. We will first establish the efficient influence function for the flip effect and then we will use it to construct multiply robust and sequentially doubly robust estimators.

5.1 Notation

To facilitate exposition, we refine our notation. First, we let

$$r_t(b_t \mid h_t) = \frac{Q_t(b_t \mid h_t)}{\mathbb{P}(A_t = b_t \mid h_t)} \tag{7}$$

be the ratio of the interventional propensity score and the true propensity score and let $r_0 = 1$ and $Q_{T+1}(A_{T+1} \mid H_{T+1}) = 1$. Then, we let $m_{T+1} = Y$, $m_T(b_T, H_T) = \mathbb{E}(Y \mid A_T = b_T, H_T)$, and recursively define

$$m_t(b_t, h_t) = \mathbb{E}\left\{ \sum_{b_{t+1}} m_{t+1}(b_{t+1}, H_{t+1}) Q_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t = b_t, H_t = h_t \right\} \tag{8}$$

as the sequential regression function for $t < T$.

5.2 Efficient influence function

The identification result in Theorem 1 suggests a "plug-in estimator" by plugging estimates of the relevant nuisance functions into each of the relevant formulas and then taking a sample average. With well-specified parametric models for the nuisance functions, the plug-in estimator can achieve $\sqrt{n}$-convergence rates. However, if the models are mis-specified, the plug-in estimator can be biased [Kang and Schafer, 2007, Vansteelandt et al., 2012]. Meanwhile, if the nuisance functions are estimated with nonparametric methods, the plug-in estimator will typically inherit slower-than-$\sqrt{n}$ nonparametric convergence rates. This motivates estimators based on nonparametric efficiency theory [Bickel et al., 1993, Tsiatis, 2006, van der Vaart, 2000].

The first-order bias of the nonparametric plug-in can be characterized by the efficient influence function of the functional, which can be thought of as its first derivative in a von Mises expansion [von Mises, 1947]. The efficient influence function can be used to construct estimators that can achieve $\sqrt{n}$-convergence with nonparametric estimators for the nuisance functions. The next result establishes the efficient influence function of $\mathbb{E}\{Y(\overline{D}_T)\}$.

Proposition 4. *Let ψ denote an identified flip effect $\mathbb{E}\{Y(\overline{D}_T)\}$ from Theorem 1 with smooth weight function. Moreover, let*

$$\begin{aligned}\phi_t(b_t; A_t, H_t) = &\Big\{2\mathbb{1}(b_t = a_t) - 1\Big\}\Big\{\mathbb{1}(A_t = a_t) - \mathbb{P}(A_t = a_t \mid H_t)\Big\} \\ &\cdot \Big[1 - s_t\{\mathbb{P}(A_t = a_t \mid H_t)\} + s_t'\{\mathbb{P}(A_t = a_t \mid H_t)\}\Big\{1 - \mathbb{P}(A_t = a_t \mid H_t)\Big\}\Big]\end{aligned}$$

where $s_t'(y) = \frac{\partial}{\partial x} s_t(x)\big|_{x=y}$. Further suppose that the outcome Y has bounded variance and the weight function is constructed such that $r_t(A_t \mid H_t)$ is bounded. Then, the centered efficient influence function of ψ under a nonparametric model is

$$\begin{aligned}&\varphi(Z) = \varphi_m(Z) + \varphi_Q(Z) \text{ where} \\ &\varphi_m(Z) = \sum_{t=0}^{T} \left\{\prod_{s=0}^{t} r_s(A_s \mid H_s)\right\} \left\{\sum_{b_{t+1}} m_{t+1}(b_{t+1}, H_{t+1}) Q_{t+1}(b_{t+1} \mid H_{t+1}) - m_t(A_t, H_t)\right\}, \\ &\varphi_Q(Z) = \sum_{t=1}^{T} \left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} m_t(b_t, H_t) \phi_t(b_t; A_t, H_t).\end{aligned}$$

The efficient influence function in Proposition 4 follows a typical structure: $\varphi(Z)$ consists of a plug-in estimator minus the true functional, plus weighted residual terms. The first component, $\varphi_m(Z)$, represents the efficient influence function that would arise if $Q_t(A_t \mid H_t)$ were known and did not require estimation. The second component, $\varphi_Q(Z)$, emerges from the necessity of estimating this quantity. It includes $\phi_t(b_t; A_t, H_t)$, which is the centered efficient influence function of $\mathbb{E}\{Q_t(A_t = b_t \mid H_t)\}$.

The result requires bounded variance of $\varphi(Z)$, which is guaranteed if Y has bounded variance and $r_t(A_t \mid H_t)$ is bounded for all $t \leq T$. The boundedness condition on r_t can be guaranteed through appropriate construction of the weight function. All the smooth weights in Table 2 except for the smooth trimming weights satisfy it immediately. For smooth trimming weights, the bound can be satisfied by construction. For instance, choosing $s(x) = 1 - \exp(-kx)$ ensures $r_t(A_t \mid H_t) = \frac{Q_t(A_t|H_t)}{\mathbb{P}(A_t|H_t)}$ is bounded since $s(x)/x \leq k$.

5.3 Multiply robust-style estimator

The efficient influence function in Proposition 4 suggests a multiply robust-style estimator.

Algorithm 1 (Multiply robust-style estimator)**.** *Randomly split the data into K folds, denoted by $\{\mathbf{Z}_k\}_{k=1}^{K}$.*
For $k = 1$ to K:

1. *Let $\cup_{l \neq k} \mathbf{Z}_l$ be the training data and $\mathbf{Z}_k$ be the evaluation data.*

2. *In the training data, regress A_t on H_t and obtain propensity score models $\widehat{\mathbb{P}}_{-k}(A_t \mid H_t)$.*

3. *In the evaluation data, compute the interventional propensity scores $\widehat{Q}_k(A_t \mid H_t)$, ratios $\widehat{r}_{t,k}(A_t \mid H_t)$, and efficient influence functions $\widehat{\phi}_{t,k}(b_t; A_t, H_t)$ for all subjects and timepoints using $\widehat{\mathbb{P}}_{-k}(A_t \mid H_t)$.*

For $t = T$ to $t = 1$:

1. *For $k = 1$ to K:*

   (a) *If $t = T$, then $\widehat{P}_{T+1}(H_{T+1}) = Y$. Otherwise, pseudo-outcome $\widehat{P}_{t+1}(H_{t+1})$ is available from the previous step in loop (see step #2 below).*

   (b) *Regress $\widehat{P}_{t+1}(H_{t+1})$ against A_t and H_t in the training data to obtain models $\widehat{m}_{t,-k}(A_t, H_t)$.*

   (c) *In the evaluation data, obtain predictions $\widehat{m}_{t,k}(0, H_t), \widehat{m}_{t,k}(1, H_t)$.*

2. *Across the full data, compute pseudo-outcomes $\widehat{P}_t(H_t) = \widehat{m}_t(0, H_t)\widehat{Q}_t(0 \mid H_t) + \widehat{m}_t(1, H_t)\widehat{Q}_t(1 \mid H_t)$ to use in the next step.*

Then,

1. *Compute the plug-in estimate $\widehat{m}_0 = \mathbb{P}_n\{\widehat{P}_1(X_1)\}$ using the last pseudo-outcome from the prior loop.*

2. *For all subjects in the data, compute the centered efficient influence function values as*

$$
\begin{aligned}
\widehat{\varphi}(Z) = &\sum_{t=0}^{T} \left\{ \prod_{s=0}^{t} \widehat{r}_s(A_s \mid H_s) \right\} \left\{ \sum_{b_{t+1}} \widehat{m}_{t+1}(b_{t+1}, H_{t+1})\widehat{Q}_{t+1}(b_{t+1} \mid H_{t+1}) - \widehat{m}_t(A_t, H_t) \right\} \\
&+ \sum_{t=1}^{T} \left\{ \prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s) \right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\widehat{\phi}_t(b_t; A_t, H_t)
\end{aligned}
$$

where

$$
\begin{aligned}
\widehat{\phi}_t(b_t; A_t, H_t) = &\Big\{ 2\mathbb{1}(b_t = a_t) - 1 \Big\} \Big\{ \mathbb{1}(A_t = a_t) - \widehat{\mathbb{P}}(A_t = a_t \mid H_t) \Big\} \\
&\cdot \Big[1 - s_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t)\} + s_t'\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t)\}\big\{1 - \widehat{\mathbb{P}}(A_t = a_t \mid H_t)\big\} \Big]
\end{aligned}
$$

Finally, output the point estimate and variance estimates

$$
\begin{aligned}
\widehat{\psi} &:= \widehat{m}_0 + \mathbb{P}_n\{\widehat{\varphi}(Z)\} \text{ and} \\
\widehat{\sigma}^2 &:= \mathbb{P}_n\big\{\widehat{\varphi}(Z)^2\big\}
\end{aligned}
$$

Algorithm 1 constructs an estimate of the efficient influence function by first estimating $\{\widehat{Q}_t\}_{t=1}^T$ and then working sequentially from $t = T$ to $t = 1$ to estimate $\{\widehat{m}_t\}_{t=1}^T$. This sequential regression formulation is the same as in Kennedy [2019], and uses an estimated pseudo-outcome $\widehat{P}_{t+1}(H_{t+1})$ in a regression to estimate $m_t(A_t, H_t)$. An alternative is the targeted maximum likelihood estimator (TMLE) in Díaz et al. [2023], which offers the same asymptotic guarantees. Algorithm 1 also employs sample splitting and cross-fitting to avoid relying on Donsker or other complexity conditions on the nuisance function estimators [Chen et al., 2022, Chernozhukov et al., 2018, Robins et al., 2008, van der Vaart and Wellner, 1996, Zheng and van der Laan, 2010]. Therefore, we are agnostic to the choice of regression method.

5.3.1 Multiply robust-style convergence guarantees

The next result provides the primary convergence guarantee for this estimator: a bound on its bias. We then show that the estimator satisfies a rate multiply robust-style result, in the sense of Rotnitzky et al. [2021], describing when $\sqrt{n}$-efficiency and asymptotic normality hold.

Theorem 2. *Under the setup of Proposition 4, let $\widehat{\psi}$ denote a point estimate from Algorithm 1 and let*

- $\widetilde{m}_t(A_t, H_t) = \mathbb{E}\left\{\sum_{b_{t+1}} \widehat{m}_{t+1}(b_{t+1}, H_{t+1})\widehat{Q}_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t, H_t\right\}$ *and*
- $\widehat{\pi}_t(H_t) := \widehat{\mathbb{P}}(A_t = 1 \mid H_t)$.

Suppose $\exists\ C < \infty$ *such that* $\mathbb{P}\{\widehat{m}_t(A_t, H_t) \leq C\} = \mathbb{P}\{m_t(A_t, H_t) \leq C\} = 1$ *for* $t \leq T$. *Then,*

$$\left|\mathbb{E}\left(\widehat{\psi} - \psi\right)\right| \lesssim \min\Bigg\{ \sum_{t=1}^T \|\widehat{\pi}_t - \pi_t\| \|\widehat{m}_t - m_t\| + \|\widehat{\pi}_t - \pi_t\|^2, \\ \sum_{t=1}^T \|\widehat{m}_t - \widetilde{m}_t\| \Big(\sum_{s=1}^t \|\widehat{\pi}_s - \pi_s\| \Big) + \|\widehat{\pi}_t - \pi_t\| \Big(\sum_{s=1}^t \|\widehat{\pi}_s - \pi_s\| \Big) \Bigg\}.$$

Theorem 2 provides a bound on the bias of the multiply robust-style estimator. Under the assumptions of Proposition 4, we only require that both the true and estimated regression functions m_t and $\widehat{m}_t$ are bounded. This result provides three new contributions in longitudinal data and sheds light on estimating WATEs in single-timepoint data:

1. **Simultaneous bounds on the bias.** We establish that two bounds hold at once, so the bias can be bounded by their minimum. To our knowledge, this is novel. This arises because the total bias decomposes into a sum of errors from $t = 1$ to $t = T$, with the first part of the minimum obtained by decomposing the error at

future timepoints via the sequential regression $\widehat{m}_t$, and the second by decomposing the error at past timepoints via $\{\widehat{Q}_s\}_{s\leq t}$.

2. **Extension of Díaz et al. [2023, Theorem 3].** The first part of the minimum extends Díaz et al. [2023, Theorem 3] to a one-step estimator and to dynamic stochastic interventions with unknown Q_t. In this setting, the dependence on future timepoints $s \geq t$ is contained in $\|\widehat{m}_t - m_t\|$, which captures the errors from sequential regressions from $s = T$ to $s = t$, as well as from the propensity scores $\{\widehat{Q}_s\}_{s>t}$ that define the pseudo-outcomes. Our bound is new in explicitly incorporating $\|\widehat{\pi}_t - \pi_t\|^2$, reflecting the fact that the flipping probabilities must be estimated.

3. **A tighter bound than Kennedy [2019, Theorem 3].** The second part of our bound depends only on the sequential regression error at time t, ignoring pseudo-outcome estimation. Specifically, it involves $\|\widehat{m}_t - \widetilde{m}_t\|$, whereas Kennedy [2019, Theorem 3] upper bounds the same term by $\|\widehat{m}_t - m_t\|$, which implicitly includes additional error from future propensity scores and sequential regressions (as discussed in point 2.).

4. **Doubly robust-style bounds for WATE estimation when $T = 1$.** For single-timepoint data, this estimator yields doubly robust-style bounds for estimating the class of WATEs described in Section 2. In some cases, this was already known; for example, it is well-established that one can upper bound the bias of an estimator for the ATO by $\sum_{a\in\{0,1\}} \|\widehat{\mu}_a - \mu_a\|\|\widehat{\pi} - \pi\| + \|\widehat{\pi} - \pi\|^2$, where $\mu_a = \mathbb{E}(Y \mid A = a, X)$, because the ATO can be re-written as $\frac{\mathbb{E}\{\text{cov}(A,Y|X)\}}{\mathbb{E}\{\mathbb{V}(A|X)\}}$. Our result includes that bias bound as a special case. For more complex smooth weight functions, such as Shannon's entropy weights—which take the form $f(X) = -\big[\pi(X)\log\pi(X) + \{1-\pi(X)\}\log\{1-\pi(X)\}\big]$—the literature has not, to our knowledge, developed doubly robust-style estimators that allow for nonparametric nuisance estimation while accounting for uncertainty in propensity score estimation during weight construction. Our result confirms this is possible.

This bound on the bias indicates when weak convergence is possible.

Corollary 1 (Multiple robustness and weak convergence)**.** *Under the setup of Theorem 2, let $\widehat{\sigma}^2$ be a variance estimate from Algorithm 1. Suppose $\|\widehat{\varphi} - \varphi\| = o_{\mathbb{P}}(1)$ and*

$$\min\Bigg\{ \sum_{t=1}^{T} \|\widehat{\pi}_t - \pi_t\|\|\widehat{m}_t - m_t\| + \|\widehat{\pi}_t - \pi_t\|^2, \\ \sum_{t=1}^{T} \|\widehat{m}_t - \widetilde{m}_t\|\Big(\sum_{s=1}^{t}\|\widehat{\pi}_s - \pi_s\|\Big) + \|\widehat{\pi}_t - \pi_t\|\Big(\sum_{s=1}^{t}\|\widehat{\pi}_s - \pi_s\|\Big)\Bigg\} = o_{\mathbb{P}}(n^{-1/2}).$$

Then,

$$\sqrt{\frac{n}{\widehat{\sigma}^2}}\left(\widehat{\psi}-\psi\right) \rightsquigarrow N(0,1). \tag{9}$$

Corollary 1 provides a multiply robust-style guarantee, showing conditions under which the estimator achieves root-n convergence to a Gaussian limit. Specifically, the first requirement ensures that the estimated efficient influence function is consistent, and the second is the crucial multiply robust-style bound on the bias. In particular, the product of the nuisance estimation errors from Theorem 2 must converge to zero at a rate of $n^{-1/2}$. This condition is achievable under nonparametric assumptions on the nuisance functions (e.g., smoothness, sparsity, or bounded variation), where each nuisance function can be estimated at a $n^{-1/4}$ rate [Györfi et al., 2002].

Remark 4. When Q_t is unknown, a *model* multiply robust-style result for consistency (in the sense of Rotnitzky et al. [2021]) is less immediately interesting than in settings with known Q_t. When Q_t is unknown, consistent estimation of $\{\pi_t\}_{t=1}^T$ is necessary, but is also sufficient, to guarantee $\widehat{\psi} \xrightarrow{\mathbb{P}} \psi$. However, Theorem 2 implies a new result when Q_t is known: $2(T+1)$ model multiple robustness; see, e.g., Díaz et al. [2023, Lemma 2] for details on $T+1$ model multiple robustness. In other words, our result implies that the typical multiply robust estimator was twice as robust as was previously realized.

5.4 Sequentially doubly robust-style estimator

The multiply robust-style estimator can be improved to a sequentially doubly robust-style estimator. One can gain intuition for how this is possible by examining the estimated pseudo-outcome $\widehat{P}_{t+1}(H_{t+1})$ in Algorithm 1. Notice that regressing $\widehat{P}_{t+1}(H_{t+1}) = \widehat{m}_{t+1}(0, H_{t+1})\widehat{Q}_t(0 \mid H_{t+1}) + \widehat{m}_{t+1}(1, H_{t+1})\widehat{Q}_{t+1}(1 \mid H_{t+1})$ against $\{A_t, H_t\}$ corresponds to using a plug-in estimator for $m_t(A_t, H_t)$. This estimator can be improved by debiasing this pseudo-outcome. For sequential regressions with longitudinal data, this was first observed in Luedtke et al. [2017] and Rotnitzky et al. [2017], and recently extended to LMTPs in Díaz et al. [2023]. This general approach — debiasing a pseudo-outcome — has also been applied to conditional effect estimation, continuous dose-response curve estimation, and censoring [Kennedy, 2023, Kennedy et al., 2017, McClean et al., 2024a, Rubin and van der Laan, 2007]. An adaptation of the estimator in Algorithm 1 is inspired by the following lemma.

Lemma 1. *Under the setup of Proposition 4, define* $Y = m_{T+1} = \sum_{b_{T+1}} m_{T+1}\left(Q_{T+1} + \phi_{T+1}\right)$ *and recursively define for* $t = T$ *to* $t = 1$

$$P_t^*(Z) = \sum_{b_t} m_t(b_t, H_t)\Big\{Q_t(b_t \mid H_t) + \phi_t(b_t; A_t, H_t)\Big\}$$
$$+ \sum_{s=t}^{T}\left\{\prod_{k=t}^{s} r_k(A_k \mid H_k)\right\}\left\{\sum_{b_{s+1}} m_{s+1}(b_{s+1}, H_{s+1})\Big\{Q_{s+1}(b_{s+1} \mid H_{s+1}) + \phi_{s+1}(b_{s+1}; A_{s+1}, H_{s+1})\Big\} - m_s(A_s, H_s)\right\}.$$

Then,

$$\mathbb{E}\left\{P^*_{t+1}(Z) \mid A_t, H_t\right\} = m_t(A_t, H_t). \tag{10}$$

Moreover, suppose access to fixed nuisance estimates $\left\{\widehat{m}^*_s, \widehat{Q}_s\right\}_{s=t+1}^{T}$ *to construct* $\widehat{P}^*_{t+1}(Z)$*. Then,*

$$\begin{aligned}
&\mathbb{E}\left\{\widehat{P}^*_{t+1}(Z) - m_t(A_t, H_t) \mid A_t, H_t\right\} = \\
&\quad \sum_{s=t+1}^{T} \mathbb{E}\left[\left\{\prod_{k=t+1}^{s-1} \widehat{r}_k(A_k \mid H_k)\right\}\left\{m_s(A_s, H_s) - \widehat{m}^*_s(A_s, H_s)\right\}\left\{\widehat{r}_s(A_s \mid H_s) - r_s(A_s \mid H_s)\right\} \mid A_t, H_t\right] \\
&+ \sum_{s=t+1}^{T} \mathbb{E}\left[\left\{\prod_{k=t+1}^{s-1} \widehat{r}_k(A_k \mid H_k)\right\} \sum_{b_s} \widehat{m}^*_s(A_s, H_s)\left\{\widehat{Q}_s(b_s \mid H_s) + \widehat{\phi}_s(b_s; A_s, H_s) - Q_s(b_s \mid H_s)\right\} \mid A_t, H_t\right].
\end{aligned} \tag{11}$$

This lemma proposes the debiased pseudo-outcome, P^*_t, then shows that it is indeed unbiased (in (10)) and that its error, if it were estimated, is a product of errors (in (11)). This mirrors Lemma 1 in Díaz et al. [2023], Lemma 1 in Luedtke et al. [2017], and Lemma 2 in Rotnitzky et al. [2017]. However, this result is new because it accounts for the error in estimating the interventional propensity score Q_t, from which the second term in the bias decomposition in (11) arises. This result inspires a new, sequentially doubly robust-style estimator, which amends Algorithm 1.

Algorithm 2 (Sequentially doubly robust-style estimator)**.** *Use Algorithm 1 with the following amendments to the sequential regression loop:*

- *In step 1(a), let* $\widehat{P}^*_{T+1}(Z) = Y$*.*
- *In step 1(b), regress* $\widehat{P}^*_{t+1}(Z)$ *against* A_t *and* H_t *and label these models* $\widehat{m}^*_{t,-k}(A_t, H_t)$*.*
- *In step 1(c), label the predictions in the evaluation data* $\widehat{m}^*_{t,k}(0, H_t), \widehat{m}^*_{t,k}(1, H_t)$*.*
- *In step 2, when constructing pseudo-outcomes, use the transformation* $\widehat{P}^*_t(Z)$ *which uses available nuisance estimates* $\left\{\widehat{m}^*_s, \widehat{Q}_s\right\}_{s=t+1}^{T}$*.*

Finally, construct a point estimate and variance estimate as

$$\widehat{\psi}^* := \widehat{m}^*_0 = \mathbb{P}_n\{\widehat{P}^*_1(Z)\} \text{ and}$$
$$\widehat{\sigma}^2 := \frac{n}{n-1}\mathbb{P}_n\left[\left\{\widehat{P}^*_1(Z) - \widehat{m}^*_0\right\}^2\right].$$

The estimator is similar to the multiply robust estimator in Algorithm 1, but uses the debiased pseudo-outcomes and debiased sequential regression estimates. A consequence of this is that $\widehat{P}^*_1(Z)$ already takes the same form as the un-centered efficient influence function from Proposition 4 and the point estimate and variance estimate can be constructed using $\widehat{P}^*_1(Z)$, rather than constructing an estimate of the efficient influence function.

5.4.1 Sequentially doubly robust-style convergence guarantees

The next result gives the sequentially doubly robust-style properties of the estimator.

Theorem 3. *Under the setup of Theorem 2, let $\widehat{\psi}^*$ denote a point estimate from Algorithm 2 and let $\widetilde{m}_t^*(A_t, H_t) = \mathbb{E}\left\{\widehat{P}_{t+1}^*(Z) \mid A_t, H_t\right\}$. Moreover, suppose $\exists\, C < \infty$ such that $\mathbb{P}\{\widehat{m}_t^*(A_t, H_t) \leq C\} = 1$ for $t \leq T$. Then,*

$$\left|\mathbb{E}\left(\widehat{\psi}^* - \psi\right)\right| \lesssim \sum_{t=1}^{T} \|\widehat{\pi}_t - \pi_t\| \Big(\|\widehat{m}_t^* - \widetilde{m}_t^*\| + \|\widehat{\pi}_t - \pi_t\|\Big).$$

Theorem 3 shows that the estimate is sequentially doubly robust-style, in that its bias can decomposed as a sum of errors across timepoints where the error at each timepoint only depends on the propensity score at that timepoint and the sequential regression estimate at that timepoint. Note that $\|\widehat{m}_t^* - \widetilde{m}_t^*\|$ only captures the error from the sequential regression at t; there is no dependence on $s > t$ through the pseudo-outcome. Therefore, we have the following asymptotic convergence guarantee.

Corollary 2. *Under the setup of Theorem 3, let $\widehat{\sigma}^2$ be a variance estimate from Algorithm 2. Suppose $\|\widehat{P}_1^* - P_1^*\| = o_{\mathbb{P}}(1)$ and*

$$\sum_{t=1}^{T} \|\widehat{\pi}_t - \pi_t\| \Big(\|\widehat{m}_t^* - \widetilde{m}_t^*\| + \|\widehat{\pi}_t - \pi_t\|\Big) = o_{\mathbb{P}}(n^{-1/2}).$$

Then,

$$\sqrt{\frac{n}{\widehat{\sigma}^2}}\big(\widehat{\psi} - \psi\big) \rightsquigarrow N(0,1). \tag{12}$$

Corollary 2 provides a sequentially doubly robust-style guarantee for weak convergence. It improves on Corollary 1 because it only requires the nuisance estimators converge at a rate of $n^{-1/2}$ in product *at each timepoint*. There is no dependence across timepoints, unlike in Corollary 1.

6 Ongoing work

In ongoing work, we are developing software implementation of our multiply robust and sequentially doubly robust estimators in R. Future versions of this manuscript will include data analyses and simulations with these estimators.

References

Ryan M Andrews and Vanessa Didelez. Insights into the cross-world independence assumption of causal mediation analysis. *Epidemiology*, 32(2):209–219, 2021.

Peter J Bickel, Chris AJ Klaassen, Y Ritov, and JA Wellner. *Efficient and adaptive estimation for semiparametric models*, volume 4. Springer, 1993.

Matteo Bonvini, Alec McClean, Zach Branson, and Edward H Kennedy. Incremental causal effects: an introduction and review. In *Handbook of matching and weighting adjustments for causal inference*, pages 349–372. Chapman and Hall/CRC, 2023.

Zach Branson, Edward H Kennedy, Sivaraman Balakrishnan, and Larry Wasserman. Causal effect estimation after propensity score trimming with continuous treatments. *arXiv preprint arXiv:2309.00706*, 2023.

Qizhao Chen, Vasilis Syrgkanis, and Morgane Austern. Debiased machine learning without sample-splitting for stable estimators. *Advances in Neural Information Processing Systems*, 35:3096–3109, 2022.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 2018.

Eric R Cohn, Eli Ben-Michael, Avi Feller, and José R Zubizarreta. Balancing weights for causal inference. In *Handbook of Matching and Weighting Adjustments for Causal Inference*, pages 293–312. Chapman and Hall/CRC, 2023.

Richard K Crump, V Joseph Hotz, Guido W Imbens, and Oscar A Mitnik. Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1):187–199, 2009.

Iván Díaz and Mark van der Laan. Population intervention causal effects based on stochastic interventions. *Biometrics*, 68(2):541–549, 2012.

Iván Díaz, Nicholas Williams, Katherine L Hoffman, and Edward J Schenck. Nonparametric causal effects based on longitudinal modified treatment policies. *Journal of the American Statistical Association*, 118(542):846–857, 2023.

László Györfi, Michael Kohler, Adam Krzyzak, Harro Walk, et al. *A distribution-free theory of nonparametric regression*, volume 1. Springer, 2002.

Jens Hainmueller. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, 20(1):25–46, 2012.

Sebastian Haneuse and Andrea Rotnitzky. Estimation of the effect of interventions that modify the received treatment. *Statistics in medicine*, 32(30):5260–5277, 2013.

Miguel Hernán and James Robins. *Causal Inference: What if.* Boca Raton: Chapman & Hall/CRC, 2020.

Aksel KG Jensen, Theis Lange, Olav L Schjørring, and Maya L Petersen. Identification and responses to positivity violations in longitudinal studies: an illustration based on invasively mechanically ventilated icu patients. *Biostatistics & Epidemiology*, 8(1):e2347709, 2024.

Joseph DY Kang and Joseph L Schafer. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science*, 22(4):523–539, 2007.

Edward Kennedy, Sivaraman Balakrishnan, and Larry Wasserman. Semiparametric counterfactual density estimation. *Biometrika*, 110(4):875–896, 2023.

Edward H Kennedy. Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*, 114(526):645–656, 2019.

Edward H Kennedy. Towards optimal doubly robust estimation of heterogeneous causal effects. *Electronic Journal of Statistics*, 17(2):3008–3049, 2023.

Edward H Kennedy, Zongming Ma, Matthew D McHugh, and Dylan S Small. Nonparametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(4):1229–1245, 2017.

Edward H Kennedy, Sivaraman Balakrishnan, and Max G'Sell. Sharp instruments for classifying compliers and generalizing causal effects. *The Annals of Statistics*, 48(4): 2008–2030, 2020.

Samir Khan and Johan Ugander. Doubly-robust and heteroscedasticity-aware sample trimming for causal inference. *arXiv preprint arXiv:2210.10171*, 2022.

Alexander W Levis, Edward H Kennedy, Alec McClean, Sivaraman Balakrishnan, and Larry Wasserman. Stochastic interventions, sensitivity analysis, and optimal transport. *arXiv preprint arXiv:2411.14285*, 2024.

Fan Li and Fan Li. Propensity score weighting for causal inference with multiple treatments. *The Annals of Applied Statistics*, 13(4):2389–2415, 2019.

Fan Li, Kari Lock Morgan, and Alan M Zaslavsky. Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400, 2018.

Alexander R Luedtke, Oleg Sofrygin, Mark J van der Laan, and Marco Carone. Sequential double robustness in right-censored longitudinal models. *arXiv preprint arXiv:1705.02459*, 2017.

Alec McClean, Zach Branson, and Edward H Kennedy. Nonparametric estimation of conditional incremental effects. *Journal of Causal Inference*, 12(1):20230024, 2024a.

Alec McClean, Yiting Li, Sunjae Bae, Mara A McAdams-DeMarco, Iván Díaz, and Wenbo Wu. Fair comparisons of causal parameters with many treatments and positivity violations. *arXiv preprint arXiv:2410.13522*, 2024b.

Kelly L Moore, Romain Neugebauer, Mark J van der Laan, and Ira B Tager. Causal inference in epidemiological studies with strong confounding. *Statistics in medicine*, 31 (13):1380–1404, 2012.

Judea Pearl. *Causality*. Cambridge University Press, 2009.

Maya L Petersen, Kristin E Porter, Susan Gruber, Yue Wang, and Mark J Van Der Laan. Diagnosing and responding to violations in the positivity assumption. *Statistical methods in medical research*, 21(1):31–54, 2012.

Thomas S Richardson and James M Robins. Single world intervention graphs (swigs): A unification of the counterfactual and graphical approaches to causality. *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper*, 128 (30):2013, 2013.

James Robins. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical modelling*, 7(9-12):1393–1512, 1986.

James Robins, Lingling Li, Eric Tchetgen Tchetgen, and Aad van der Vaart. Higher order influence functions and minimax estimation of nonlinear functionals. In *Institute of Mathematical Statistics Collections*, pages 335–421. Institute of Mathematical Statistics, 2008.

James M Robins, Miguel A Hernán, and Uwe Siebert. Effects of multiple interventions. *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*, 1:2191–2230, 2004.

Andrea Rotnitzky, James Robins, and Lucia Babino. On the multiply robust estimation of the mean of the g-functional. *arXiv preprint arXiv:1705.08582*, 2017.

Andrea Rotnitzky, Ezequiel Smucler, and James M Robins. Characterization of parameters with a mixed bias property. *Biometrika*, 108(1):231–238, 2021.

Daniel Rubin and Mark J van der Laan. A doubly robust censoring unbiased transformation. *The international journal of biostatistics*, 3(1), 2007.

Kyle Schindl, Shuying Shen, and Edward H Kennedy. Incremental effects for continuous exposures. *arXiv preprint arXiv:2409.11967*, 2024.

Mats J Stensrud, JD Laurendeau, and Aaron L Sarvet. Optimal regimes for algorithm-assisted human decision-making. *Biometrika*, page asae016, 2024.

James H Stock. Nonparametric policy analysis. *Journal of the American Statistical Association*, 84(406):567–575, 1989.

Sarah L Taubman, James M Robins, Murray A Mittleman, and Miguel A Hernán. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *International journal of epidemiology*, 38(6):1599–1611, 2009.

Eric J Tchetgen Tchetgen and Tyler J VanderWeele. On causal inference in the presence of interference. *Statistical methods in medical research*, 21(1):55–75, 2012.

Anastasios A Tsiatis. *Semiparametric theory and missing data*, volume 4. Springer, 2006.

Aad W van der Vaart. *Asymptotic statistics*, volume 3. Cambridge University Press, 2000.

Aad W van der Vaart and Jon A Wellner. *Weak convergence and empirical processes*. Springer, 1996.

Stijn Vansteelandt, Maarten Bekaert, and Gerda Claeskens. On model selection and model misspecification in causal inference. *Statistical methods in medical research*, 21(1):7–30, 2012.

Richard von Mises. On the asymptotic distribution of differentiable statistical functions. *The annals of mathematical statistics*, 18(3):309–348, 1947.

Shu Yang and Peng Ding. Asymptotic inference of causal effects with observational studies trimmed by the estimated propensity scores. *Biometrika*, 105(2):487–493, 2018.

Jessica G Young, Miguel A Hernán, and James M Robins. Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. *Epidemiologic methods*, 3(1):1–19, 2014.

Shuxi Zeng, Fan Li, and Liangyuan Hu. Propensity score weighting analysis of survival outcomes using pseudo-observations. *Statistica Sinica*, 33(3):2161–2184, 2023.

Wenjing Zheng and Mark J van der Laan. Asymptotic theory for cross-validated targeted maximum likelihood estimation. *U.C. Berkeley Division of Biostatistics Working Paper Series*, 2010.

Xiang Zhou and Aleksei Opacic. Marginal interventional effects. *arXiv preprint arXiv:2206.10717*, 2022.

Appendix

This appendix does the following:

Appendix A provides efficient estimators for the average treatment value.

Appendix B provides proofs for the results in Section 2.

Appendix C provides proofs for the results in Section 4.

Appendix D provides proofs for the results in Section 5.

Appendix E provides proofs for the results in Appendix A.

A Estimating the average treatment value

For longitudinal interventional flip interventions defined in (3), we must also estimate $\mathbb{E}(D_t)$ and $\mathbb{E}(D'_t)$. These estimands are identified in Theorem 1. In this section, we focus on estimating $\mathbb{E}(D_T)$ when $\{D_1, \ldots, D_T\}$ are flip interventions with smooth weights, as in Section 5. Adapting the estimator for $t < T$ is straightforward. Using the notation from the main paper, the identified estimand is

$$\psi_D = \mathbb{E}\left[Q_T(A_T = 1 \mid H_T)\left\{\prod_{t=1}^{T-1} r_t(A_t \mid H_t)\right\}\right] \tag{13}$$

where $r_t(A_t \mid H_t) = \frac{Q_t(A_t|H_t)}{\mathbb{P}(A_t|H_t)}$,

$$Q_t(A_t = b_t \mid H_t) = \mathbb{P}(A_T = b_T \mid H_T) + \{2\mathbb{1}(b_t = a_t) - 1\}\, s_t\{\mathbb{P}(A_t = a_t \mid H_t)\}\{1 - \mathbb{P}(A_t = a_t \mid H_t)\},$$

and s_t is a smooth and twice differentiable function of the propensity score with bounded derivatives.

We define the sequential regression at time T as $m_T(0, H_T) = 0$ and $m_T(1, H_T) = 1$ and then recursively define

$$m_t(A_t, H_t) = \mathbb{E}\left\{\sum_{b_{t+1}} m_{t+1}(b_{t+1}, H_{t+1}) Q_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t, H_t\right\}.$$

for $t < T$. Then, we can derive the efficient influence function.

Proposition 5. *Let ψ_D be as in (13). Under the setup of Proposition 4 the centered efficient influence function of ψ_D under a nonparametric model is*

$$\varphi_D(Z) = \sum_{t=1}^{T}\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\}\left[\sum_{b_t} m_t(b_t, H_t)\left\{Q_t(b_t \mid H_t) + \phi_t(b_t; A_t, H_t)\right\} - m_{t-1}(A_{t-1}, H_{t-1})\right].$$

Proof. The proof is deferred to the Appendix E. □

One can construct a multiply robust-style estimator based on this efficient influence function. It is a straightforward adaptation of Algorithm 1 from the main paper.

Algorithm 3 (Multiply robust-style estimator)**.** *Randomly split the data into K folds, denoted by $\{\mathbf{Z}_k\}_{k=1}^K$.*
For $k = 1$ to K:

1. *Let $\cup_{l \neq k} \mathbf{Z}_l$ be the training data and $\mathbf{Z}_k$ be the evaluation data.*
2. *In the training data, regress A_t on H_t and obtain propensity score models $\widehat{\mathbb{P}}_{-k}(A_t \mid H_t)$.*
3. *In the evaluation data, compute the interventional propensity scores $\widehat{Q}_k(A_t \mid H_t)$, ratios $\widehat{r}_{t,k}(A_t \mid H_t)$, and efficient influence functions $\widehat{\phi}_{t,k}(b_t; A_t, H_t)$ for all subjects and timepoints using $\widehat{\mathbb{P}}_{-k}(A_t \mid H_t)$.*

For $t = T - 1$ to $t = 1$:

1. *For $k = 1$ to K:*
   - *(a) If $t = T - 1$, then $\widehat{P}_T(H_T) = \widehat{Q}_T(A_T = 1 \mid H_T)$. Otherwise, pseudo-outcome $\widehat{P}_{t+1}(H_{t+1})$ is available from the previous step in loop (see step #2 below).*
   - *(b) Regress $\widehat{P}_{t+1}(H_{t+1})$ against A_t and H_t in the training data to obtain models $\widehat{m}_{t,-k}(A_t, H_t)$.*
   - *(c) In the evaluation data, obtain predictions $\widehat{m}_{t,k}(0, H_t), \widehat{m}_{t,k}(1, H_t)$.*
2. *Across the full data, compute pseudo-outcomes $\widehat{P}_t(H_t) = \widehat{m}_t(0, H_t)\widehat{Q}_t(0 \mid H_t) + \widehat{m}_t(1, H_t)\widehat{Q}_t(1 \mid H_t)$ to use in the next step.*

Then,

1. *Compute the plug-in estimate $\widehat{m}_0 = \mathbb{P}_n\{\widehat{P}_1(X_1)\}$ using the last pseudo-outcome from the prior loop.*
2. *For all subjects in the data, compute the centered efficient influence function values as*

$$\widehat{\varphi}_D(Z) = \sum_{t=1}^T \left\{ \prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s) \right\} \left[\sum_{b_t} \widehat{m}_t(b_t, H_t) \left\{ \widehat{Q}_t(b_t \mid H_t) + \widehat{\phi}_t(b_t; A_t, H_t) \right\} - \widehat{m}_{t-1}(A_{t-1}, H_{t-1}) \right]$$

where

$$\begin{aligned} \widehat{\phi}_t(b_t; A_t, H_t) = &\left\{ 2\mathbb{1}(b_t = a_t) - 1 \right\} \left\{ \mathbb{1}(A_t = a_t) - \widehat{\mathbb{P}}(A_t = a_t \mid H_t) \right\} \\ &\cdot \left[1 - s_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t)\} + s_t'\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t)\}\{1 - \widehat{\mathbb{P}}(A_t = a_t \mid H_t)\} \right]. \end{aligned}$$

Finally, output the point estimate and variance estimates

$$\widehat{\psi}_D := \widehat{m}_0 + \mathbb{P}_n\{\widehat{\varphi}_D(Z)\} \text{ and}$$
$$\widehat{\sigma}_D^2 := \mathbb{P}_n\{\widehat{\varphi}_D(Z)^2\}.$$

We have the following bound on the bias of the estimator from Algorithm 3.

Theorem 4. *Under the setup of Theorem 2, let ψ_D be as in (13) and $\widehat{\psi}_D$ denote the estimator from Algorithm 3. Moreover, let $\widetilde{m}_t(A_t, H_t) = \mathbb{E}\left\{ \sum_{b_{t+1}} \widehat{m}_{t+1}(b_{t+1}, H_{t+1})\widehat{Q}_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t, H_t \right\}$. Then,*

$$\left| \mathbb{E}\left(\widehat{\psi}_D - \psi_D \right) \right| \lesssim \min\left\{ \sum_{t=1}^T \|\widehat{m}_t - m_t\| \|\widehat{\pi}_t - \pi_t\| + \|\widehat{\pi}_t - \pi_t\|^2, \right.$$

$$\sum_{t=1}^{T} \|\widehat{m}_t - \widetilde{m}_t\| \left(\sum_{s=1}^{t} \|\widehat{\pi}_s - \pi_s\| \right) + \|\widehat{\pi}_t - \pi_t\| \left(\sum_{s=1}^{t} \|\widehat{\pi}_s - \pi_s\| \right).$$

Proof. This follows by the same analysis as for the efficient estimator of $\mathbb{E}\{Y(\overline{D}_T)\}$ and using the algebra in the proof of Proposition 5. We omit it for brevity. □

Then, we have the following convergence guarantee.

Corollary 3. *Under the setup of Theorem 4, let $\widehat{\sigma}_D^2$ be a variance estimate from Algorithm 3. Suppose $\|\widehat{\varphi}_D - \varphi_D\| = o_{\mathbb{P}}(1)$ and*

$$\min\left\{ \sum_{t=1}^{T} \|\widehat{m}_t - m_t\| \|\widehat{\pi}_t - \pi_t\| + \|\widehat{\pi}_t - \pi_t\|^2, \right.$$
$$\left. \sum_{t=1}^{T} \|\widehat{m}_t - \widetilde{m}_t\| \left(\sum_{s=1}^{t} \|\widehat{\pi}_s - \pi_s\| \right) + \|\widehat{\pi}_t - \pi_t\| \left(\sum_{s=1}^{t} \|\widehat{\pi}_s - \pi_s\| \right) \right\} = o_{\mathbb{P}}(n^{-1/2}).$$

Then,

$$\sqrt{\frac{n}{\widehat{\sigma}_D^2}} (\widehat{\psi}_D - \psi_D) \rightsquigarrow N(0,1).$$

As in the main paper with the following result, we motivate the sequentially doubly robust-style estimator with a result on the debiased pseudo-outcome.

Lemma 2. *Under the setup of Proposition 5, define $P_T^*(Z) = Q_T(A_t = 1 \mid H_T) + \phi_T(1; A_T, H_T)$ and recursively define for $t = T-1$ to $t = 1$*

$$P_t^*(Z) = \sum_{b_t} m_t(b_t, H_t) \Big\{ Q_t(b_t \mid H_t) + \phi_t(b_t; A_t, H_t) \Big\}$$
$$+ \sum_{s=t}^{T-1} \left\{ \prod_{k=t}^{s} r_k(A_k \mid H_k) \right\} \left\{ \sum_{b_{s+1}} m_{s+1}(b_{s+1}, H_{s+1}) \Big\{ Q_{s+1}(b_{s+1} \mid H_{s+1}) + \phi_{s+1}(b_{s+1}; A_{s+1}, H_{s+1}) \Big\} - m_s(A_s, H_s) \right\}.$$

Then,

$$\mathbb{E}\left\{ P_{t+1}^*(Z) \mid A_t, H_t \right\} = m_t(A_t, H_t). \tag{14}$$

Moreover, suppose access to fixed nuisance estimates $\left\{ \widehat{m}_s^*, \widehat{Q}_s \right\}_{s=t+1}^{T}$ to construct $\widehat{P}_{t+1}^*(Z)$. Then,

$$\mathbb{E}\left\{ \widehat{P}_{t+1}^*(Z) - m_t(A_t, H_t) \mid A_t, H_t \right\} =$$
$$\sum_{s=t+1}^{T} \mathbb{E}\left[\left\{ \prod_{k=t+1}^{s-1} \widehat{r}_k(A_k \mid H_k) \right\} \Big\{ m_s(A_s, H_s) - \widehat{m}_s^*(A_s, H_s) \Big\} \Big\{ \widehat{r}_s(A_s \mid H_s) - r_s(A_s \mid H_s) \Big\} \mid A_t, H_t \right]$$
$$+ \sum_{s=t+1}^{T} \mathbb{E}\left[\left\{ \prod_{k=t+1}^{s-1} \widehat{r}_k(A_k \mid H_k) \right\} \sum_{b_s} \widehat{m}_s^*(A_s, H_s) \left\{ \widehat{Q}_s(b_s \mid H_s) + \widehat{\phi}_s(b_s; A_s, H_s) - Q_s(b_s \mid H_s) \right\} \mid A_t, H_t \right]. \tag{15}$$

Proof. See Appendix E. □

This result suggests the simple amendment to the estimator from Algorithm 3.

Algorithm 4 (Sequentially doubly robust-style estimator)**.** *Use Algorithm 3 with the following amendments to the sequential regression loop:*

- *In step 1(a), let* $\widehat{P}_T^*(Z) = \widehat{Q}_T(A_T = 1 \mid H_T) + \widehat{\phi}_T(1; A_T, H_T)$.
- *In step 1(b), regress* $\widehat{P}_{t+1}^*(Z)$ *against* A_t *and* H_t *and label these models* $\widehat{m}_{t,-k}^*(A_t, H_t)$.
- *In step 1(c), label the predictions in the evaluation data* $\widehat{m}_{t,k}^*(0, H_t), \widehat{m}_{t,k}^*(1, H_t)$.
- *In step 2, when constructing pseudo-outcomes, use the transformation* $\widehat{P}_t^*(Z)$ *which uses available nuisance estimates* $\left\{\widehat{m}_s^*, \widehat{Q}_s\right\}_{s=t+1}^T$.

Finally, construct a point estimate and variance estimate as

$$\widehat{\psi}_D^* := \widehat{m}_0^* = \mathbb{P}_n\{\widehat{P}_1^*(Z)\} \text{ and}$$
$$\widehat{\sigma}_D^2 := \frac{n}{n-1}\mathbb{P}_n\left[\left\{\widehat{P}_1^*(Z) - \widehat{m}_0^*\right\}^2\right].$$

Finally, we have the following bias bound and convergence guarantee.

Theorem 5. *Under the setup of Theorem 4, let* $\widehat{\psi}_D^*$ *denote a point estimate from Algorithm 4 and let* $\widetilde{m}_t^*(A_t, H_t) = \mathbb{E}\left\{\widehat{P}_{t+1}^*(Z) \mid A_t, H_t\right\}$. *Then,*

$$\left|\mathbb{E}\left(\widehat{\psi}_D^* - \psi_D\right)\right| \lesssim \sum_{t=1}^T \|\widehat{\pi}_t - \pi_t\|\Big(\|\widehat{m}_t^* - \widetilde{m}_t^*\| + \|\widehat{\pi}_t - \pi_t\|\Big).$$

B Proofs for Section 2: Proposition 1

Proof. The numerator satisfies

$$\begin{aligned}
&\mathbb{E}\big[Y\{D_f(1)\} - Y\{D_f(0)\}\big] = \mathbb{E}\left(\mathbb{E}\big[Y\{D_f(1)\} - Y\{D_f(0)\} \mid X\big]\right) \\
&= \sum_{a\in\{0,1\}} \mathbb{E}\Big[\mathbb{E}\{Y(a) \mid D_f(1) = a, X\}\mathbb{P}\{D_f(1) = a \mid X\}\Big] - \mathbb{E}\Big[\mathbb{E}\{Y(a) \mid D_f(0) = a, X\}\mathbb{P}\{D_f(0) = a \mid X\}\Big] \\
&= \sum_{a\in\{0,1\}} \mathbb{E}\left[\mathbb{E}\{Y(a) \mid X\}\mathbb{P}\{D_f(1) = a \mid X\}\right] - \mathbb{E}\left[\mathbb{E}\{Y(a) \mid X\}\mathbb{P}\{D_f(0) = a \mid X\}\right] \\
&= \mathbb{E}\left(\mathbb{E}\{Y(1) - Y(0) \mid X\}\Big[\mathbb{P}\{D_f(1) = 1 \mid X\} - \mathbb{P}\{D_f(0) = 1 \mid X\}\Big]\right)
\end{aligned}$$

where the first line follows by iterated expectations and the second by taking the expectations over $D_f(1)$ and $D_f(0)$. The third follows by the unconfoundedness assumption because $D_f(a)$ is only a function of A, conditional on X, so that $Y(a) \perp\!\!\!\perp D_f(a) \mid X$. The final line follows by rearranging. Next, notice that

$$\mathbb{P}\{D_f(1) = 1 \mid X\} = \mathbb{P}\Big[\mathbb{1}(A = 1)A + \mathbb{1}(A = 0) \cdot \mathbb{1}\{V \le f(X)\} = 1 \mid X\Big] = \pi(X) + \{1 - \pi(X)\}f(X)$$

and

$$\mathbb{P}\{D_f(0) = 1 \mid X\} = \mathbb{P}\left[\mathbb{1}(A = 1) \cdot \mathbb{1}\{V > f(X)\} = 1 \mid X\right] = \pi(X)\{1 - f(X)\}.$$

The difference is

$$\mathbb{P}\{D_f(1) = 1 \mid X\} - \mathbb{P}\{D_f(0) = 1 \mid X\} = f(X).$$

Therefore, the numerator satisfies

$$\mathbb{E}\big[Y\{D_f(1)\} - Y\{D_f(0)\}\big] = \mathbb{E}\left[\mathbb{E}\{Y(1) - Y(0) \mid X\}f(X)\right].$$

By the same argument, the denominator satisfies

$$\mathbb{E}\big\{D_f(1) - D_f(0)\big\} = \mathbb{E}\left\{f(X)\right\}.$$

The result follows. □

C Proofs for Section 4

C.1 Helper results

For the identification results, we provide several helper lemmas. We also slightly amend our notation from the main paper, so we can specify the dependence of random variables on the counterfactual past interventions. In what follows, we make the following "assumption" (which corresponds to the setup in the main paper).

Assumption 3. Suppose the following setup:

- $\{D_1, D_2(D_1), \dots, D_T(\overline{D}_{T-1})\}$ denote a set of treatment decisions where $D_t(\overline{D}_{t-1}) = f_t\{A_t(\overline{D}_{t-1}), H_t(\overline{D}_{t-1}), V_t\}$ for some deterministic function f_t, where $V_1, \dots, V_T$ are mutually independent and $V_t \perp\!\!\!\perp Z$ for all $t \in \{1, \dots, T\}$,
- $\overline{X}_t(\overline{a}_{t-1})$ denotes the natural covariate history under an intervention that sets treatment to $\overline{a}_{t-1}$ up until time $t-1$,
- $H_t(\overline{a}_{t-1}) = \big\{\overline{X}_{t-1}(\overline{a}_{t-1}), \overline{D}_{t-1} = \overline{a}_{t-1}\big\}$ denotes the natural covariate history and the intervention treatment history,
- $A_t(\overline{a}_{t-1})$ denotes the natural value of treatment after history $H_t(\overline{a}_{t-1})$,
- $D_t(\overline{a}_{t-1}) = d_t\{A_t(\overline{a}_{t-1}), H_t(\overline{a}_{t-1}), V_t\}$, i.e., the treatment decision at time t is random via $A_t(\overline{a}_{t-1}), H_t(\overline{a}_{t-1})$, and V_t,
- $Y(\overline{a}_t, \underline{D}_{t+1})$ denotes the potential outcome under an intervention that sets treatment $\overline{a}_t$ up until time t and assigns treatment according to treatment decisions $\underline{D}_{t+1} = \{D_{t+1}(\overline{a}_t), D_{t+2}(D_{t+1}, \overline{a}_t), \dots D_T(\overline{a}_t, D_{t+1}, \dots, D_{T-1})\}$ thereafter.

The first result gives us the important consistency steps we use within the g-formula. In essence, it says that if we condition on an observed history up until the end of timepoint $t-1$, then the counterfactual history in timepoint t is equal to the observed history. An important corollary is that, conditional on an observed history up until the end of timepoint $t-1$, the propensity scores at timepoint t are identified.

Proposition 6. *Conditional on* $\{\overline{X}_{t-1}, \overline{A}_{t-1} = \overline{b}_{t-1}\}$,

- $H_t(\overline{b}_{t-1}) = \big\{\overline{X}_t, \overline{A}_{t-1} = \overline{b}_{t-1}\big\}$ *and*
- $A_t(\overline{b}_{t-1}) = A_t$.

As a consequence, conditional on $\{\overline{X}_{t-1}, \overline{A}_{t-1} = \bar{b}_{t-1}\}$,

$$\mathbb{P}\{A_t(\bar{b}_{t-1}) = b_t \mid H_t(\bar{b}_{t-1})\} = \mathbb{P}(A_t = b_t \mid \overline{X}_t, \overline{A}_{t-1} = \bar{b}_{t-1}).$$

In other words, the propensity score at timepoint t *is identified.*

Proof. These follow by the consistency assumption in the NPSEM. □

The next result gives us the crucial exchangeability we need when the intervention depends on the natural value of treatment.

Lemma 3. *Under Assumption 2 and Assumption 3,*

$$\begin{aligned} A_t(\overline{a}_{t-1}) \perp\!\!\!\perp Y(\overline{a}_t, \underline{D}_{t+1}) \mid H_t(\overline{a}_{t-1}) \text{ and} \\ D_t(\overline{a}_{t-1}) \perp\!\!\!\perp Y(\overline{a}_t, \underline{D}_{t+1}) \mid H_t(\overline{a}_{t-1}). \end{aligned}$$

Proof. Conditional on $H_t(\overline{a}_{t-1})$, $A_t(\overline{a}_{t-1})$ only depends on the random variable $U_{A,t}$. Meanwhile, $Y(\overline{a}_t, \underline{D}_{t+1})$ depends on $(\underline{V}_{t+1}, \underline{U}_{A,t+1}, \underline{U}_{X,t+1}, U_Y)$. Both results then follow by Assumption 2 and the assumption on $\overline{V}_T$. □

The next result gives us the same exchangeability result when the intervention does not depend on the natural value of treatment. It only requires standard exchangeability.

Lemma 4. *Under Assumption 3, suppose instead that*

$$D_t(\overline{a}_{t-1}) = d_t\{\mathbb{P}\{A_t(\overline{a}_{t-1}) \mid H_t(\overline{a}_{t-1})\}, H_t(\overline{a}_{t-1}), V_t\}$$

for some function f_t*; i.e., the intervention only depends on the natural propensity score, the natural covariate history, and auxiliary randomness, but not the natural value of treatment. Then, under only Assumption 1,*

$$\begin{aligned} A_t(\overline{a}_{t-1}) \perp\!\!\!\perp Y(\overline{a}_t, \underline{D}_{t+1}) \mid H_t(\overline{a}_{t-1}) \text{ and} \\ D_t(\overline{a}_{t-1}) \perp\!\!\!\perp Y(\overline{a}_t, \underline{D}_{t+1}) \mid H_t(\overline{a}_{t-1}). \end{aligned}$$

Proof. Conditional on the natural covariate history, the intervention is only random via V_t. Meanwhile, $\underline{D}_{t+1}$ are only random via the future natural covariate history and random variables $\underline{V}_{t+1}$. Therefore, conditional on $H_t(\overline{a}_{t-1})$, $Y(\overline{a}_t, \underline{D}_{t+1})$ is only random via $(\underline{U}_{X,t+1}, U_Y, \underline{V}_{t+1})$. Therefore, the result follows by Assumption 1 and the assumption on $\overline{V}_T$. □

Remark 5. The construction in Lemma 4 applies to our suggested stochastic amendments in Section 4.1, such as

$$\begin{aligned} D_t = \mathbb{1}\bigg(V_t \leq \mathbb{1}(a_t = 1) f_t^c\{a_t; H_t(\overline{D}_{t-1})\} \\ + \Big[1 - f_t^c\{a_t; H_t(\overline{D}_{t-1})\}\Big] \mathbb{P}\{A_t(\overline{D}_{t-1} = 1 \mid H_t(\overline{D}_{t-1})\} \bigg) \end{aligned}$$

for flip interventions.

Finally, we establish a result for positivity, which shows that the flip interventions avoid positivity violations under the condition in Theorem 1.

Lemma 5. *Under the setup of Theorem 1,* $\mathbb{P}\big\{\mathbb{P}(A_t = b_t \mid H_t) = 0 \implies Q_t(b_t \mid H_t) = 0\big\} = 1$ *for* $b_t \in \{0, 1\}$ *for both flip interventions satisfying* $\mathbb{P}(A_t = a_t \mid H_t) \implies f_t(a_t; H_t) = 0$*, where* f_t *is defined in Theorem 1.*

Proof. For flip interventions, when a_t is the target treatment

$$Q_t(a_t \mid H_t) = \mathbb{P}(A_t = a_t \mid H_t) + \{1 - \mathbb{P}(A_t = 1 - a_t \mid H_t)\}f_t(a_t; H_t).$$

By construction, $\mathbb{P}(A_t = a_t \mid H_t) = 0 \implies f_t(a_t; H_t) = 0$; therefore, $\mathbb{P}(A_t = a_t \mid H_t) = 0$ implies

$$Q_t(a_t \mid H_t) = 0 + 1 \cdot 0 = 0.$$

Meanwhile, $\mathbb{P}(A_t = 1 - a_t \mid H_t) = 0$ implies

$$Q_t(a_t \mid H_t) = 1 + 0 = 1$$

which itself implies $Q_t(1 - a_t \mid H_t) = 0$. □

C.2 Proof of Theorem 1

Finally, we have the full proof of the main theorem.

Proof. First, we have

$$\begin{aligned}
\mathbb{E}\big\{Y(\overline{D}_T)\big\} &= \mathbb{E}\Big[\mathbb{E}\big\{Y(\overline{D}_T) \mid X_1\big\}\Big] = \mathbb{E}\Big[\mathbb{E}\big\{Y(\overline{D}_T) \mid D_1, X_1\big\} \mid X_1\Big] \\
&= \mathbb{E}\Big[\sum_{b_1} \mathbb{E}\big\{Y(b_1, \underline{D}_2) \mid D_1 = b_1, X_1\big\}\mathbb{P}(D_1 = b_1 \mid X_1)\Big] \\
&= \mathbb{E}\Big[\sum_{b_1} \mathbb{E}\big\{Y(b_1, \underline{D}_2) \mid A_1 = b_1, X_1\big\}Q_1(b_1 \mid X_1)\Big] \\
&\equiv \sum_{b_1} \int_{\mathcal{X}_1} \mathbb{E}\big\{Y(b_1, \underline{D}_2) \mid A_1 = b_1, x_1\big\}Q_1(b_1 \mid x_1)d\mathbb{P}(x_1)
\end{aligned}$$

where the first line follows by iterated expectations on X_1 and then on X_1 and D_1, the second by taking the expectation over D_1, the third by Lemma 3 in the inner expectation and by the definition of D_1 in the outer probability and Proposition 6 to identify the counterfactual trimming indicator, and the fourth by linearity of expectation and definition. Note that, by Lemma 5, the outer expectation is well-defined. The inner expectation might not be well-defined, but $Q_1(b_1 \mid x_1) = 0$ whenever that occurs.

The rest of the proof will follow by induction. We address the $t = 2$ step. We have

$$\begin{aligned}
\mathbb{E}\big\{Y(b_1, \underline{D}_2) \mid A_1 = b_1, X_1\big\} &= \mathbb{E}\Big[\mathbb{E}\big\{Y(b_1, \underline{D}_2) \mid X_2(b_1), A_1 = b_1, X_1\big\} \mid A_1 = b_1, X_1\Big] \\
&\equiv \mathbb{E}\Big[\mathbb{E}\big\{Y(b_1, \underline{D}_2) \mid H_2(b_1)\big\} \mid A_1 = b_1, X_1\Big] \\
&= \mathbb{E}\Big[\sum_{b_2} \mathbb{E}\big\{Y(b_1, b_2, \underline{D}_3) \mid D_2(b_1) = b_2, H_2(b_1)\big\}\mathbb{P}\{D_2(b_1) = b_2 \mid H_2(b_1)\} \mid A_1 = b_1, X_1\Big] \\
&= \mathbb{E}\Big[\sum_{b_2} \mathbb{E}\big\{Y(b_1, b_2, \underline{D}_3) \mid A_2(b_1) = b_2, H_2(b_1)\big\}Q_2(b_2 \mid b_1, \overline{X}_2) \mid A_1 = b_1, X_1\Big]
\end{aligned}$$

$$= \sum_{b_2} \mathbb{E}\Big[\mathbb{E}\big\{Y(b_1, b_2, \underline{D}_3) \mid A_2 = b_2, X_2, A_1 = b_1, X_1\big\} Q_2(b_2 \mid b_1, \overline{X}_2) \mid A_1 = b_1, X_1\Big]$$

where the first line follows by iterated expectations on $X_2(b_1), A_1 = b_1, X_1$, the second line by Proposition 6, and the third by iterated expectations on $D_2(b_2), H_2(b_1)$ and then taking the expectation over $D_2(b_2)$. The fourth follows by Lemma 3 inside the expectation; meanwhile, the probability Q_2 is identified by Proposition 6. The final line follows again by Proposition 6. Again, note that by Lemma 5, the outer expectation is well-defined. The inner expectation might not be well-defined, but $Q_2(b_2 \mid H_2) = 0$ whenever that occurs.

Repeating this process $t - 2$ more times yields

$$\mathbb{E}\left\{Y(\overline{D}_T)\right\} = \sum_{\bar{b}_T \in \{0,1\}^T} \int_{\overline{\mathcal{X}}_T} \mathbb{E}\{Y(\bar{b}_t) \mid \overline{A}_T = \bar{b}_T, \overline{X}_T = \overline{x}_T\} \prod_{t=1}^{T} Q_t(b_t \mid \bar{b}_{t-1}, \overline{x}_t) d\mathbb{P}(x_t \mid \bar{b}_{t-1}, \overline{x}_{t-1}).$$

The final result follows by the consistency assumption embedded in the NPSEM. The IPW result follows by taking the expectation over A_t.

Meanwhile, the identification of $\mathbb{E}(D_t)$ follows by essentially the same argument. We'll repeat the first step here:

$$\mathbb{E}(D_t) = \sum_{b_1} \mathbb{E}\{\mathbb{E}(D_t \mid D_1 = b_1, X_1)\mathbb{P}(D_1 = b_1 \mid X_1)\} = \sum_{b_1} \mathbb{E}\{\mathbb{E}(D_t \mid A_1 = b_1, X_1) Q_1(b_1 \mid X_1)\}$$

where the first equation follows by iterated expectations, the second by Lemma 3 to exchange $D_1 = b_1$ with $A_1 = b_1$ and by the definition of D_1 to yield $\mathbb{P}(D_1 = b_1 \mid X_1) = Q_1(b_1 \mid X_1)$. The result follows by repeating this process $t - 1$ more times. □

Remark 6. Suppose the flip intervention was stochastic as in our suggested stochastic amendment at the end of Section 4.1:

$$D_t = \mathbb{1}\Bigg(V_t \leq \mathbb{1}(a_t = 1) f_t^c\{a_t; H_t(\overline{D}_{t-1})\} + \Big[1 - f_t^c\{a_t; H_t(\overline{D}_{t-1})\}\Big]\mathbb{P}\{A_t(\overline{D}_{t-1} = 1 \mid H_t(\overline{D}_{t-1})\}\Bigg).$$

Then, the sequential exchangeability step would follow by Lemma 4, which only requires standard sequential randomization in Assumption 1. Meanwhile, identification of $\mathbb{E}(D_t)$ would only require the NPSEM assumption and possibly positivity depending on the weight function, but no exchangeability assumption.

C.3 Proposition 3

Proof. We have

$$\mathbb{E}\big\{Y(\overline{D}_T)\big\} = \mathbb{E}\Big[\sum_{\bar{b}_T} \mathbb{E}\{Y(\bar{b}_T) \mid \overline{D}_T = \bar{b}_T, \overline{A}_T, \Pi_T\}\mathbb{P}(\overline{D}_T = \bar{b}_T \mid \overline{A}_T, \Pi_T)\Big]$$
$$= \mathbb{E}\Big[\sum_{\bar{b}_T} \mathbb{E}\{Y(\bar{b}_T) \mid \overline{A}_T, \Pi_T\}\mathbb{P}(\overline{D}_T = \bar{b}_T \mid \overline{A}_T, \Pi_T)\Big]$$

where the first line follows by iterated expectations on $\overline{A}_T, \Pi_T$ and then on $\overline{D}_T, \overline{A}_T, \Pi_T$ and taking the expectation over $\overline{D}_T$, and the second line follows because $\overline{D}_T \perp\!\!\!\perp Y(\bar{b}_T) \mid \overline{A}_T, \Pi_T$ because $\overline{D}_T$ is only random via V conditional on $\overline{A}_T, \Pi_T$.

The same holds for $\mathbb{E}\{Y(\overline{D}'_T)\}$. Then,

$$\mathbb{E}\{Y(\overline{D}_T) - Y(\overline{D}'_T)\} = \sum_{\bar{b}_T} \mathbb{E}\left[\mathbb{E}\{Y(\bar{b}_T) \mid \overline{A}_T, \overline{X}_T\}\{\mathbb{P}(\overline{D}_T = \bar{b}_T \mid \overline{A}_T, \Pi_T) - \mathbb{P}(\overline{D}'_T = \bar{b}_T \mid \overline{A}_T, \Pi_T)\}\right].$$

For the propensity scores, we have

$$\begin{aligned}
\mathbb{P}(\overline{D}_T = \bar{a}_T \mid \overline{A}_T, \Pi_T) &= \mathbb{1}(\overline{A}_T = \bar{a}_T) + \mathbb{1}(\overline{A}_T \neq \bar{a}_T)\Pi_T, \\
\mathbb{P}(\overline{D}_T = \bar{b}_T \mid \overline{A}_T, \Pi_T) &= \mathbb{1}(\overline{A}_T = \bar{b}_T)(1 - \Pi_T) \text{ for } \bar{b}_T \neq \bar{a}_T, \\
\mathbb{P}(\overline{D}'_T = \bar{a}'_T \mid \overline{A}_T, \Pi_T) &= \mathbb{1}(\overline{A}_T = \bar{a}'_T) + \mathbb{1}(\overline{A}_T \neq \bar{a}'_T)\Pi_T, \text{ and} \\
\mathbb{P}(\overline{D}_T = \bar{b}_T \mid \overline{A}_T, \Pi_T) &= \mathbb{1}(\overline{A}_T = \bar{b}_T)(1 - \Pi_T) \text{ for } \bar{b}_T \neq \bar{a}'_T.
\end{aligned}$$

Plugging these results into the prior display yields

$$\mathbb{E}\{Y(\overline{D}_T) - Y(\overline{D}'_T)\} = \mathbb{E}\left(\left[\mathbb{E}\{Y(\bar{a}_T) \mid \overline{A}_T, \Pi_T\} - \mathbb{E}\{Y(\bar{a}_T) \mid \overline{A}_T, \Pi_T\}\right]\Pi_T\right).$$

The result follows by iterated expectations on $\overline{A}_T, \Pi_T$. The same argument holds for the denominator of (6). □

D Proofs for Section 5

D.1 Helper lemmas of efficient influence function of Q_t

We begin with several general helper lemmas.

Lemma 6. *Under the setup of Proposition 4, $\varphi_m(Z)$ and $\varphi_Q(Z)$ are mean-zero.*

Proof. This follows by iterated expectations on H_t. □

The next two lemmas are about the efficient influence function of $\mathbb{E}\{Q_t(b_t \mid H_t)\}$ and its estimator as constructed in the body of the paper.

Lemma 7. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\phi_t(b_t; A_t, H_t) \mid H_t\} = 0.$$

and

$$\begin{aligned}
\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\} = &\left\{2\mathbb{1}(b_t = a_t) - 1\right\}\left\{\mathbb{P}(A_t = a_t) - \widehat{\mathbb{P}}(A_t = a_t \mid H_t)\right\} \\
&\cdot \left[1 - s_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t)\} + s'_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t)\}\{1 - \widehat{\mathbb{P}}(A_t = a_t \mid H_t)\}\right].
\end{aligned}$$

Proof. These follow by iterated expectations on H_t. □

In the next lemma we omit arguments for brevity, so that $\mathbb{P}_t = \mathbb{P}(A_t = a_t \mid H_t)$ and $s_t = s_t\{\mathbb{P}(A_t = a_t \mid H_t)\}$ and $\widehat{\mathbb{P}}_t$ and $\widehat{s}_t(a_t)$ are defined similarly.

Lemma 8. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) + \widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t) \mid H_t\} =$$
$$= \Big\{2\mathbb{1}(b_t = a_t) - 1\Big\} \left(\left[\frac{\widehat{s}_t''}{2} \left(\widehat{\mathbb{P}}_t - \mathbb{P}_t\right)^2 + o\{(\widehat{\mathbb{P}}_t - \mathbb{P}_t)^2\} \right] \left(\widehat{\mathbb{P}}_t - 1\right) + (\widehat{s} - s) \left(\widehat{\mathbb{P}}_t - \mathbb{P}_t\right) \right)$$

Proof. First, note that by definition $Q_t(b_t \mid H_t) = \mathbb{P}(A_t = b_t \mid H_t) + \Big\{2\mathbb{1}(b_t = a_t) - 1\Big\} s_t\{\mathbb{P}(A_t = a_t \mid H_t)\{1 - \mathbb{P}(A_t = a_t \mid H_t)\}$ and therefore

$$\mathbb{E}\left\{\widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t) \mid H_t\right\} = \Big\{2\mathbb{1}(b_t = a_t) - 1\Big\} \left\{\widehat{\mathbb{P}}_t - \mathbb{P}_t + \widehat{s}_t \cdot (1 - \widehat{\mathbb{P}}_t) - s_t \cdot (1 - \mathbb{P}_t)\right\},$$

where we omit arguments on the right-hand side. Therefore, by iterated expectations and rearranging, we have

$$\begin{aligned}
&\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) + \widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t) \mid H_t\} \\
&= \Big\{2\mathbb{1}(b_t = a_t) - 1\Big\} \left[(\mathbb{P}_t - \widehat{\mathbb{P}}_t) \left\{1 - \widehat{s}_t + \widehat{s}_t' \cdot (1 - \widehat{\mathbb{P}}_t)\right\} + \widehat{s}_t \cdot (1 - \widehat{\mathbb{P}}_t) - s_t \cdot (1 - \mathbb{P}_t) + \widehat{\mathbb{P}}_t - \mathbb{P}_t\right] \\
&= \Big\{2\mathbb{1}(b_t = a_t) - 1\Big\} \left[\widehat{s}_t' \cdot (\mathbb{P}_t - \widehat{\mathbb{P}}_t) + \widehat{s}_t - s_t - \widehat{\mathbb{P}}_t \left\{\widehat{s}_t' \cdot (\mathbb{P}_t - \widehat{\mathbb{P}}_t) + \widehat{s}_t - s_t\right\} - s_t \cdot \widehat{\mathbb{P}}_t - \widehat{s}_t \cdot (\mathbb{P}_t - \widehat{\mathbb{P}}_t) + s_t \cdot \mathbb{P}_t\right] \\
&= \Big\{2\mathbb{1}(b_t = a_t) - 1\Big\} \left[(1 - \widehat{\mathbb{P}}_t) \left\{\widehat{s}_t' \cdot (\mathbb{P}_t - \widehat{\mathbb{P}}_t) + \widehat{s}_t - s_t\right\} + (\widehat{s}_t - s_t)(\widehat{\mathbb{P}}_t - \mathbb{P}_t)\right]
\end{aligned}$$

A second-order Taylor expansion of $s_t\{\mathbb{P}(A_t = a_t \mid H_t)\}$ yields the result. Specifically,

$$(1 - \widehat{\mathbb{P}}_t) \left\{\widehat{s}_t' \cdot (\mathbb{P}_t - \widehat{\mathbb{P}}_t) + \widehat{s}_t - s_t\right\} = (\widehat{\mathbb{P}}_t - 1) \left[\frac{\widehat{s}_t''}{2} \left(\widehat{\mathbb{P}}_t - \mathbb{P}_t\right)^2 + o\left\{(\widehat{\mathbb{P}}_t - \mathbb{P}_t)^2\right\}\right]$$

□

D.2 Proposition 4 and Theorem 2

Now, we turn to establishing Proposition 4 and Theorem 2. As discussed in the body of the paper, this can be established in two ways, by unwinding the error backwards-in-time or forwards-in-time. We start with lemmas for the backwards-in-time bound, which is similar to Lemmas 5 & 6 in Kennedy [2019]. We establish the result in full, for two reasons. First, for completeness. And second, our analysis yields a different bound on the bias. Then, we establish the forwards-in-time bound. This mirrors results in Díaz et al. [2023] and others, but is new because it accounts for estimating the Q_t.

In what follows, let $\widetilde{m}_t(A_t, H_t) = \mathbb{E}\left\{\sum_{b_{t+1}} \widehat{m}_{t+1}(b_{t+1}, H_{t+1})\widehat{Q}_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t, H_t\right\}$ as in the body of the paper. In other words, $\widetilde{m}_t$ is the true sequential regression function at timepoint t where all the future information is estimated.

D.2.1 Backwards-in-time lemmas

Lemma 9. *Under the setup of Proposition 4,*

$$\begin{aligned}
\mathbb{E}\{\widehat{\varphi}_m(Z)\} &= m_0 - \widehat{m}_0 \\
&\quad + \sum_{t=0}^{T} \mathbb{E}\left[\left\{\prod_{s=0}^{t} \widehat{r}_s(A_s \mid H_s) - \prod_{s=0}^{t} r_s(A_s \mid H_s)\right\} \{\widetilde{m}_t(A_t, H_t) - \widehat{m}_t(A_t, H_t)\}\right]
\end{aligned}$$

$$+\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=0}^{t-1}r_s(A_s\mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t,H_t)\{\widehat{Q}_t(b_t\mid H_t)-Q_t(b_t\mid H_t)\}\right]$$

Proof. We have

$$\mathbb{E}\{\widehat{\varphi}_m(Z)\}=\mathbb{E}\left[\sum_{t=T}^{0}\left\{\prod_{s=0}^{t}\widehat{r}_s(A_s\mid H_s)\right\}\left\{\sum_{b_{t+1}}\widehat{m}_{t+1}(b_{t+1},H_{t+1})\widehat{Q}_{t+1}(b_{t+1}\mid H_{t+1})-\widehat{m}_t(A_t,H_t)\right\}\right]$$
$$=\mathbb{E}\left[\sum_{t=T}^{0}\left\{\prod_{s=0}^{t}\widehat{r}_s(A_s\mid H_s)\right\}\{\widetilde{m}_t(A_t,H_t)-\widehat{m}_t(A_t,H_t)\}\right]$$
$$=\mathbb{E}\left[\sum_{t=T}^{0}\left\{\prod_{s=0}^{t}\widehat{r}_s(A_s\mid H_s)-\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\{\widetilde{m}_t(A_t,H_t)-\widehat{m}_t(A_t,H_t)\}\right]$$
$$+\mathbb{E}\left[\sum_{t=T}^{0}\left\{\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\{\widetilde{m}_t(A_t,H_t)-\widehat{m}_t(A_t,H_t)\}\right]$$

where the first equality follows by definition, the second by the definition of $\widetilde{m}_t(A_t,H_t)$ and iterated expectations on A_t,H_t, and the third by adding and subtracting $\prod_{s=0}^{t}r_s(A_s\mid H_s)$. On the RHS of the final equality, the first line is second-order. Focusing on the final line in the above display, notice first that the first and last summands in the overall sum can be isolated and the sum can be re-written as

$$\mathbb{E}\left[\sum_{t=T}^{0}\left\{\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\{\widetilde{m}_t(A_t,H_t)-\widehat{m}_t(A_t,H_t)\}\right]$$
$$=\mathbb{E}\left[\left\{\prod_{t=0}^{T}r_s(A_s\mid H_s)\right\}\widetilde{m}_T(A_T,H_T)\right]$$
$$+\sum_{t=T-1}^{0}\mathbb{E}\left[\left\{\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\widetilde{m}_t(A_t,H_t)-\left\{\prod_{s=0}^{t+1}r_s(A_s\mid H_s)\right\}\widehat{m}_{t+1}(A_{t+1},H_{t+1})\right]$$
$$-\widehat{m}_0$$

The first term equals ψ because $\widetilde{m}_T(A_T,H_T)=\mathbb{E}(Y\mid A_T,H_T)$ and the last term is $\widehat{m}_0$. Meanwhile, the middle term in the above display simplifies because

$$\mathbb{E}\left[\left\{\prod_{s=0}^{t+1}r_s(A_s\mid H_s)\right\}\widehat{m}_{t+1}(A_{t+1},H_{t+1})\right]$$
$$=\mathbb{E}\left[\left\{\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\mathbb{E}\left\{\sum_{b_{t+1}}\widehat{m}_{t+1}(b_{t+1},H_{t+1})Q_{t+1}(b_{t+1}\mid H_{t+1})\mid A_t,H_t\right\}\right]$$

by iterated expectations on A_t,H_t. Combining like terms and the definition of $\widetilde{m}_t$ yield

$$\sum_{t=T-1}^{0}\mathbb{E}\left[\left\{\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\widetilde{m}_t(A_t,H_t)-\left\{\prod_{s=0}^{t+1}r_s(A_s\mid H_s)\right\}\widehat{m}_{t+1}(A_{t+1},H_{t+1})\right]$$
$$=\sum_{t=T-1}^{0}\mathbb{E}\Bigg[\left\{\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\mathbb{E}\Bigg\{\sum_{b_{t+1}}\widehat{m}_{t+1}(b_{t+1},H_{t+1})\{\widehat{Q}_{t+1}(b_{t+1}\mid H_{t+1})$$

$$- Q_{t+1}(b_{t+1} \mid H_{t+1})\} \mid A_t, H_t \Big\}\Bigg]$$

$$= \sum_{t=1}^{T} \mathbb{E}\left[\left\{\prod_{s=0}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\{\widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t)\}\right]$$

where the last line follows by re-indexing the sum and iterated expectations on A_t, H_t. □

Lemma 10. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\widehat{\varphi}_Q(Z)\} = \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right]$$
$$+ \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s) - \prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right]$$
$$+ \mathbb{E}\left(\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\Big[\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\} + \widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t)\Big]\right)$$
$$+ \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t \mid H_t)\Big\{Q_t(b_t \mid H_t) - \widehat{Q}_t(b_t \mid H_t)\Big\}\right]$$

Proof. We have

$$\mathbb{E}\{\widehat{\varphi}_Q(Z)\} = \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\widehat{\phi}_t(b_t; A_t, H_t)\right]$$
$$= \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right]$$
$$= \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s) - \prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right]$$
$$+ \mathbb{E}\left(\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t)\Big[\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\} + \widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t)\Big]\right)$$
$$+ \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t \mid H_t)\Big\{Q_t(b_t \mid H_t) - \widehat{Q}_t(b_t \mid H_t)\Big\}\right]$$

where the second equality follows by adding zero several times. □

Lemma 11. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\widehat{\varphi}(Z)\} = m_0 - \widehat{m}_0$$
$$+ \sum_{t=0}^{T} \mathbb{E}\left[\left\{\prod_{s=0}^{t} \widehat{r}_s(A_s \mid H_s) - \prod_{s=0}^{t} r_s(A_s \mid H_s)\right\}\{\widetilde{m}_t(A_t, H_t) - \widehat{m}_t(A_t, H_t)\}\right]$$

$$+\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=1}^{t-1}\widehat{r}_s(A_s \mid H_s)-\prod_{s=1}^{t-1}r_s(A_s \mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t,H_t)\mathbb{E}\{\widehat{\phi}_t(b_t;A_t,H_t)\mid H_t\}\right]$$

$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=1}^{t-1}r_s(A_s \mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t,H_t)\Big[\mathbb{E}\{\widehat{\phi}_t(b_t;A_t,H_t)\mid H_t\}+\widehat{Q}_t(b_t \mid H_t)-Q_t(b_t \mid H_t)\Big]\right).$$

Proof. The final lines in the display in the previous two lemmas cancel, yielding the result. $\square$

D.2.2 Forwards-in-time lemmas

Lemma 12. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\widehat{\varphi}_m(Z)\}=m_0-\widehat{m}_0$$
$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t}\{\widehat{m}_t(b_t,H_t)-m_t(b_t,H_t)\}\,\widehat{r}_t(b_t \mid H_t)\left\{\mathbb{P}(b_t \mid H_t)-\widehat{\mathbb{P}}(b_t \mid H_t)\right\}\right]\right)$$
$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t}m_t(b_t,H_t)\left\{\widehat{Q}_t(b_t \mid H_t)-Q_t(b_t \mid H_t)\right\}\right]\right)$$

Proof. We have

$$\mathbb{E}\{\widehat{\varphi}_m(Z)\}=\mathbb{E}\left[\sum_{t=T}^{0}\left\{\prod_{s=0}^{t}\widehat{r}_s(A_s \mid H_s)\right\}\left\{\sum_{b_{t+1}}\widehat{m}_{t+1}(b_{t+1},H_{t+1})\widehat{Q}_{t+1}(b_{t+1} \mid H_{t+1})-\widehat{m}_t(A_t,H_t)\right\}\right]$$
$$=\mathbb{E}\left[\left\{\prod_{s=1}^{T}\widehat{r}_s(A_s \mid H_s)\right\}Y\right]-\widehat{m}_0$$
$$+\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s \mid H_s)\right\}\left\{\sum_{b_t}\widehat{m}_t(b_t,H_t)\widehat{Q}_t(b_t \mid H_t)-\widehat{m}_t(A_t,H_t)\widehat{r}_t(A_t \mid H_t)\right\}\right]$$
$$=\mathbb{E}\left[\left\{\prod_{s=1}^{T}\widehat{r}_s(A_s \mid H_s)\right\}Y\right]-\widehat{m}_0$$
$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t}\{\widehat{m}_t(b_t,H_t)-m_t(b_t,H_t)\}\,\widehat{r}_t(b_t \mid H_t)\left\{\mathbb{P}(b_t \mid H_t)-\widehat{\mathbb{P}}(b_t \mid H_t)\right\}\right]\right)$$
$$+\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s \mid H_s)\right\}\left\{\sum_{b_t}m_t(b_t,H_t)\widehat{Q}_t(b_t \mid H_t)-m_t(A_t,H_t)\widehat{r}_t(A_t \mid H_t)\right\}\right]$$

where the first line follows by definition, the second by rearranging the sum, and the third by adding and subtracting m_t. The second line in the final expression follows by taking iterated expectations on H_t and gather terms. We do not manipulate the second term in the final expression above any further because it appears in the final result. Focusing on the final term, we have

$$\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s \mid H_s)\right\}\left\{\sum_{b_t}m_t(b_t,H_t)\widehat{Q}_t(b_t \mid H_t)-m_t(A_t,H_t)\widehat{r}_t(A_t \mid H_t)\right\}\right]$$

$$= \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t} m_t(b_t, H_t)\left\{\widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t)\right\}\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\left\{\sum_{b_t} m_t(b_t, H_t) Q_t(b_t \mid H_t)\right\} - m_t(A_t, H_t)\widehat{r}_t(A_t \mid H_t)\right]\right)$$

$$= \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t} m_t(b_t, H_t)\left\{\widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t)\right\}\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[m_t(A_t, H_t)\left\{r_t(A_t \mid H_t) - \widehat{r}_t(A_t \mid H_t)\right\}\right]\right)$$

where the first equality follows by adding and subtracting Q_t and the second equality by iterated expectation on H_t and gathering terms, and the final line by adding and subtracting m_t. The first term in the final display appears in the result, so we manipulate them no further.

Combining the left over terms, we have

$$\sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[m_t(A_t, H_t)\left\{r_t(A_t \mid H_t) - \widehat{r}_t(A_t \mid H_t)\right\}\right]\right) + \mathbb{E}\left[\left\{\prod_{s=1}^{T} \widehat{r}_s(A_s \mid H_s)\right\} Y\right] - \widehat{m}_0$$

$$= m_0 - \sum_{t=1}^{T} \mathbb{E}\left[\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\left\{m_t(A_t, H_t) - m_{t+1}(A_{t+1}, H_{t+1}) r_{t+1}(A_{t+1} \mid H_{t+1})\right\}\widehat{r}_t(A_t \mid H_t)\right]\right) - \widehat{m}_0$$

$$= m_0 - \widehat{m}_0$$

where the first equality follows by taking the first term out of the initial sum, which equals m_0 because it is $\mathbb{E}\{m_1(A_1, H_1) r_1(A_1 \mid H_1)\}$, and adding $\mathbb{E}\left[\left\{\prod_{s=1}^{T} \widehat{r}_s(A_s \mid H_s)\right\} Y\right]$ into the sum and combining terms, and the second equality follows by iterated expectations on H_t. Combining all the algebra above yields the result. □

Lemma 13. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\widehat{\varphi}_Q(Z)\} =$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t} \{\widehat{m}_t(b_t, H_t) - m_t(b_t, H_t)\}\, \mathbb{E}\left\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t)\right\}\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t} m_t(b_t, H_t)\left\{\widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t) + \mathbb{E}\left\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t)\right\}\right\}\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\left[\sum_{b_t} m_t(b_t, H_t)\left\{Q_t(b_t \mid H_t) - \widehat{Q}_t(b_t \mid H_t)\right\}\right]\right)$$

Proof. We have

$$\mathbb{E}\{\widehat{\varphi}_Q(Z)\} = \mathbb{E}\left[\sum_{t=T}^{1}\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\}\sum_{b_t} \widehat{m}_t(b_t, H_t)\widehat{\phi}_t(b_t; A_t, H_t)\right]$$

$$= \sum_{t=1}^{T} \mathbb{E}\left[\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} \widehat{m}_t(b_t, H_t) \mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right]$$

$$= \sum_{t=1}^{T} \mathbb{E}\left[\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} \{\widehat{m}_t(b_t, H_t) - m_t(b_t, H_t)\} \mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right]$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} m_t(b_t, H_t) \left[\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\} + \widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t)\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left[\left\{\prod_{s=1}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \sum_{b_t} m_t(b_t, H_t) \left\{Q_t(b_t \mid H_t) - \widehat{Q}_t(b_t \mid H_t)\right\}\right]$$

where the second equality follows by adding zero several times. □

Lemma 14. *Under the setup of Proposition 4,*

$$\mathbb{E}\{\widehat{\varphi}_m(Z) + \widehat{\varphi}_Q(Z)\} = m_0 - \widehat{m}_0$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \left[\sum_{b_t} \{\widehat{m}_t(b_t, H_t) - m_t(b_t, H_t)\} \widehat{r}_t(b_t \mid H_t) \left\{\mathbb{P}(b_t \mid H_t) - \widehat{\mathbb{P}}(b_t \mid H_t)\right\}\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \left[\sum_{b_t} \{\widehat{m}_t(b_t, H_t) - m_t(b_t, H_t)\} \mathbb{E}\left\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\right\}\right]\right)$$

$$+ \sum_{t=1}^{T} \mathbb{E}\left(\left\{\prod_{s=0}^{t-1} \widehat{r}_s(A_s \mid H_s)\right\} \left[\sum_{b_t} m_t(b_t, H_t) \left\{\widehat{Q}_t(b_t \mid H_t) - Q_t(b_t \mid H_t) + \mathbb{E}\left\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\right\}\right\}\right]\right)$$

Proof. The final lines in the display in the previous two lemmas cancel, yielding the result. □

Proof of Proposition 4

Proof. Lemmas 11, 7, and 8 imply that $\mathbb{E}\{\widehat{\varphi}(Z) - \varphi(Z)\}$ is a second-order product of errors in nuisance functions. By the same argument, the functional satisfies a von Mises expansion with second-order remainder term. The result follows by Kennedy et al. [2023, Lemma 2], combined with the fact that $\mathbb{V}\{\varphi(Z)\}$ is bounded because the outcome has bounded variance and $\frac{Q_t(A_t|H_t)}{\mathbb{P}(A_t|H_t)}$ is bounded by assumption. □

D.3 Theorem 2

Proof. The minimum in the result will follow by taking the minimum of the two bounds we prove below.

Backwards-in-time:
The estimator is $\mathbb{P}_n\{\widehat{m}_0 + \widehat{\varphi}(Z)\}$. Because we have iid observations, the bias then satisfies

$$\mathbb{E}\left(\widehat{\psi} - \psi\right) = \widehat{m}_0 + \mathbb{E}\{\widehat{\varphi}(Z)\} - \psi \equiv \mathbb{E}\{\widehat{\varphi}(Z)\} + \widehat{m}_0 - m_0.$$

Then, by Lemma 11,

$$\mathbb{E}\left(\widehat{\psi}-\psi\right)=\sum_{t=0}^{T}\mathbb{E}\left[\left\{\prod_{s=0}^{t}\widehat{r}_s(A_s\mid H_s)-\prod_{s=0}^{t}r_s(A_s\mid H_s)\right\}\{\widetilde{m}_t(A_t,H_t)-\widehat{m}_t(A_t,H_t)\}\right]$$
$$+\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=1}^{t-1}\widehat{r}_s(A_s\mid H_s)-\prod_{s=1}^{t-1}r_s(A_s\mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t,H_t)\mathbb{E}\{\widehat{\phi}_t(b_t;A_t,H_t)\mid H_t\}\right]$$
$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=1}^{t-1}r_s(A_s\mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t,H_t)\Big[\mathbb{E}\{\widehat{\phi}_t(b_t;A_t,H_t)\mid H_t\}+\widehat{Q}_t(b_t\mid H_t)-Q_t(b_t\mid H_t)\Big]\right)$$

Note that r_t is bounded by the construction of s_t, while $\widehat{m}_t$ is bounded by assumption. Then, by Hölder's inequality, Lemmas 7 and 8, the triangle inequality, and Cauchy-Schwarz:

$$\left|\mathbb{E}\left(\widehat{\psi}-\psi\right)\right|\lesssim\sum_{t=0}^{T}\sum_{s=1}^{t}\|\widehat{r}_s-r_s\|\|\widetilde{m}_t-\widehat{m}_t\|$$
$$+\sum_{t=1}^{T}\sum_{s=1}^{t-1}\|\widehat{r}_s-r_s\|\Big(\|\widehat{\mathbb{P}}(a_t)-\mathbb{P}(a_t)\|+\|\widehat{\mathbb{P}}(b_t)-\mathbb{P}(b_t)\|\Big)$$
$$+\sum_{t=1}^{T}\Big(\|\widehat{\mathbb{P}}(a_t)-\mathbb{P}(a_t)\|^2+\|\widehat{\mathbb{P}}(b_t)-\mathbb{P}(b_t)\|^2\Big).$$

We can streamline this decomposition further, as in the statement of the result. First, note that $\widehat{r}_0=r_0=1$ by definition. Second, with binary treatment, $\sum_{a_t\in\{0,1\}}\|\widehat{\mathbb{P}}(a_t)-\mathbb{P}(a_t)\|\lesssim\|\widehat{\pi}_t-\pi_t\|$ where $\pi_t(H_t)\equiv\mathbb{P}(A_t=1\mid H_t)$. Third, $\|\widehat{r}_s-r_s\|\lesssim\|\widehat{\pi}_t-\pi_t\|$ by Taylor expansion. Then, the final line above simplifies to

$$\sum_{t=1}^{T}\sum_{s=1}^{t}\|\widehat{\pi}_s-\pi_s\|\|\widehat{m}_t-\widetilde{m}_t\|+\sum_{t=1}^{T}\sum_{s=1}^{t-1}\|\widehat{\pi}_s-\pi_s\|\|\widehat{\pi}_t-\pi_t\|+\sum_{t=1}^{T}\|\widehat{\pi}_t-\pi_t\|^2$$
$$=\sum_{t=1}^{T}\sum_{s=1}^{t}\|\widehat{\pi}_s-\pi_s\|\Big(\|\widehat{m}_t-\widetilde{m}_t\|+\|\widehat{\pi}_t-\pi_t\|\Big).$$

Forwards-in-time:
By the same argument above and Lemma 14,

$$\mathbb{E}\left(\widehat{\psi}-\psi\right)=$$
$$\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s\mid H_s)\right\}\left[\sum_{b_t}\{\widehat{m}_t(b_t,H_t)-m_t(b_t,H_t)\}\,\widehat{r}_t(b_t\mid H_t)\left\{\mathbb{P}(b_t\mid H_t)-\widehat{\mathbb{P}}(b_t\mid H_t)\right\}\right]\right)$$
$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s\mid H_s)\right\}\left[\sum_{b_t}\{\widehat{m}_t(b_t,H_t)-m_t(b_t,H_t)\}\,\mathbb{E}\left\{\widehat{\phi}_t(b_t;A_t,H_t)\mid H_t\right\}\right]\right)$$
$$+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=0}^{t-1}\widehat{r}_s(A_s\mid H_s)\right\}\left[\sum_{b_t}m_t(b_t,H_t)\left\{\widehat{Q}_t(b_t\mid H_t)-Q_t(b_t\mid H_t)+\mathbb{E}\left\{\widehat{\phi}_t(b_t;A_t,H_t)\mid H_t\right\}\right\}\right]\right)$$

Note that $\widehat{r}_t$ is bounded by the construction of s_t, while m_t is bounded by assumption. Then, by Hölder's inequality, Lemmas 7 and 8, the triangle inequality, and Cauchy-Schwarz:

$$\begin{aligned}\left|\mathbb{E}\left(\widehat{\psi}-\psi\right)\right| &\lesssim \sum_{t=1}^{T} \|\widehat{m}_t - m_t\| \|\widehat{\mathbb{P}}(b_t) - \mathbb{P}(b_t)\| \\ &+ \sum_{t=1}^{T} \|\widehat{m}_t - m_t\| \Big(\|\widehat{\mathbb{P}}(a_t) - \mathbb{P}(a_t)\| + \|\widehat{\mathbb{P}}(b_t) - \mathbb{P}(b_t)\| \Big) \\ &+ \sum_{t=1}^{T} \Big(\|\widehat{\mathbb{P}}(a_t) - \mathbb{P}(a_t)\|^2 + \|\widehat{\mathbb{P}}(b_t) - \mathbb{P}(b_t)\|^2 \Big).\end{aligned}$$

We can streamline this decomposition further, as in the statement of the result. First, with binary treatment, $\sum_{a_t \in \{0,1\}} \|\widehat{\mathbb{P}}(a_t) - \mathbb{P}(a_t)\| \lesssim \|\widehat{\pi}_t - \pi_t\|$ where $\pi_t(H_t) \equiv \mathbb{P}(A_t = 1 \mid H_t)$. Second, $\|\widehat{r}_t - r_t\| \lesssim \|\widehat{\pi}_t - \pi_t\|$ by Taylor expansion. Then, the final line above simplifies to

$$\left|\mathbb{E}\left(\widehat{\psi}-\psi\right)\right| \lesssim \sum_{t=1}^{T} \Big(\|\widehat{m}_t - m_t\| + \|\widehat{\pi}_t - \pi_t\| \Big) \|\widehat{\pi}_t - \pi_t\|.$$

□

D.4 Corollary 1

Proof. We have

$$\begin{aligned}\widehat{\psi} - \psi &= \widehat{m}_0 + \mathbb{P}_n\{\widehat{\varphi}(Z)\} - m_0 \\ &= (\mathbb{P}_n - \mathbb{P})\{\varphi(Z)\} + (\mathbb{P}_n - \mathbb{P})\{\widehat{\varphi}(Z) - \varphi(Z)\} + \widehat{m}_0 + \mathbb{P}\{\widehat{\varphi}(Z)\} - m_0 \\ &= (\mathbb{P}_n - \mathbb{P})\{\varphi(Z)\} + (\mathbb{P}_n - \mathbb{P})\{\widehat{\varphi}(Z) - \varphi(Z)\} + \mathbb{E}(\widehat{\psi} - \psi).\end{aligned}$$

where the first line follows by definition, the second by adding zero and because $\mathbb{P}\{\varphi(Z)\} = 0$, and the third line by the definition of the estimator $\widehat{\psi}$. The second term is $o_{\mathbb{P}}(n^{-1/2})$ by Chebyshev's inequality and the assumption that $\|\widehat{\varphi} - \varphi\| = o_{\mathbb{P}}(1)$ (cf. Kennedy et al. [2020, Lemma 2]). Meanwhile, the third term equals the bias term in Theorem 2. This is $o_{\mathbb{P}}(n^{-1/2})$ by assumption. Therefore,

$$\sqrt{\frac{n}{\mathbb{V}\{\varphi(Z)\}}}(\widehat{\psi} - \psi) = \sqrt{\frac{n}{\mathbb{V}\{\varphi(Z)\}}}(\mathbb{P}_n - \mathbb{P})\{\varphi(Z)\} + o_{\mathbb{P}}(1) \rightsquigarrow N(0,1)$$

by the central limit theorem and because $\mathbb{V}\{\varphi(Z)\}$ is bounded because Y has bounded variance and r_t is bounded.

Finally, note that $\widehat{\sigma}^2 \xrightarrow{p} \mathbb{V}\{\varphi(Z)\}$ because $\|\widehat{\varphi} - \varphi\| = o_{\mathbb{P}}(1)$. Therefore, the result follows by Slutsky's theorem. □

D.5 Lemma 1

Proof. The first result follows by repeated applications of iterated expectations. Notice that the residual terms are mean zero by iterated expectations, leaving only the plug-in term. Then, $\mathbb{E}\{\phi_{t+1}(b_{t+1}; A_{t+1}, H_{t+1}) \mid H_{t+1}\} = 0$ by Lemma 7. And finally, by definition,

$$\mathbb{E}\left\{\sum_{b_{t+1}} m_{t+1}(b_{t+1}, H_{t+1}) Q_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t, H_t\right\} = m_t(A_t, H_t).$$

The second result follows by induction. Throughout, we will omit arguments. Starting with the final residual, when $s = T$, we have

$$\begin{aligned}
\mathbb{E}\left\{\left(\prod_{k=t+1}^{T} \widehat{r}_k\right)(Y - \widehat{m}_T) \mid A_t, H_t\right\} &= \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)(m_T - \widehat{m}_T)\,\widehat{r}_T \mid A_t, H_t\right\} \\
&= \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)(m_T - \widehat{m}_T)\,(\widehat{r}_T - r_T) \mid A_t, H_t\right\} \\
&\quad + \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)(m_T - \widehat{m}_T)\, r_T \mid A_t, H_t\right\}
\end{aligned}$$

where the first equality follows by iterated expectations on A_T, H_T and the second by adding and subtracting r_T. The first term in the final expression appears in the result, so we manipulate it no further. The next step is the induction step.

Consider the second term in the display above and the penultimate residual, when $s = T - 1$. We have

$$\begin{aligned}
&\mathbb{E}\left\{\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)(m_T - \widehat{m}_T)\, r_T \mid A_t, H_t\right\} + \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)\left(\sum_{b_T} \widehat{m}_T(\widehat{Q}_T + \widehat{\phi}_T) - \widehat{m}_{T-1}\right) \mid A_t, H_t\right\} \\
&= \mathbb{E}\left[\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)\left\{\sum_{b_T}(m_T - \widehat{m}_T)\, Q_T + \sum_{b_T} \widehat{m}_T(\widehat{Q}_T + \widehat{\phi}_T) - \widehat{m}_{T-1}\right\} \mid A_t, H_t\right] \\
&= \mathbb{E}\left[\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)\left\{\sum_{b_T} \widehat{m}_T\left(\widehat{Q}_T + \widehat{\phi}_T - Q_T\right) + m_{T-1} - \widehat{m}_{T-1}\right\} \mid A_t, H_t\right] \\
&= \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-1} \widehat{r}_k\right)\left\{\sum_{b_T} \widehat{m}_T\left(\widehat{Q}_T + \widehat{\phi}_T - Q_T\right)\right\} \mid A_t, H_t\right] \\
&+ \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-2} \widehat{r}_k\right)(m_{T-1} - \widehat{m}_{T-1})\,(\widehat{r}_{T-1} - r_{T-1}) \mid A_t, H_t\right\} \\
&+ \mathbb{E}\left\{\left(\prod_{k=t+1}^{T-2} \widehat{r}_k\right)(m_{T-1} - \widehat{m}_{T-1})\, r_{T-1} \mid A_t, H_t\right\}
\end{aligned}$$

where the first equality follows by gathering terms, the second by iterated expectations and the definition of m_{T-1}, and the third by adding and subtracting r_{T-1}. The first and second lines in the final display appear in the statement of the result. The third line can be combined with the earlier residual, for $s = T - 2$ using the step we just outlined. This argument can be continued all the way to $s = t + 1$.

For the final step, when $s = t + 1$, we will be left with

$$\begin{aligned}
&\mathbb{E}\left[(m_{t+1} - \widehat{m}_{t+1})\, r_{t+1} + \sum_{b_{t+1}} \widehat{m}_{t+1}\left(\widehat{Q}_{t+1} + \widehat{\phi}_{t+1}\right) \mid A_t, H_t\right] - m_t(A_t, H_t) \\
&= \mathbb{E}\left[\sum_{b_{t+1}} (m_{t+1} - \widehat{m}_{t+1})\, Q_{t+1} + \sum_{b_{t+1}} \widehat{m}_{t+1}\left(\widehat{Q}_{t+1} + \widehat{\phi}_{t+1}\right) \mid A_t, H_t\right] - m_t(A_t, H_t) \\
&= \mathbb{E}\left[\sum_{b_{t+1}} \widehat{m}_{t+1}\left(\widehat{Q}_{t+1} + \widehat{\phi}_{t+1} - Q_{t+1}\right) \mid A_t, H_t\right]
\end{aligned}$$

where the first equality follows by iterated expectations and the second by canceling $\mathbb{E}\left\{\sum_{b_{t+1}} m_{t+1} Q_{t+1} \mid A_t, H_t\right\} - m_t(A_t, H_t) = 0$. □

D.6 Theorem 3

Proof. We omit arguments throughout. We have

$$\begin{aligned}
\mathbb{E}\left(\widehat{\psi}^* - \psi\right) &= \mathbb{E}\left(\widehat{m}_0^* - m_0\right) \\
&= \mathbb{E}\left(\widehat{P}_1^*(Z) - m_0\right) \\
&= \sum_{t=1}^{T} \mathbb{E}\left[\left(\prod_{k=1}^{t-1} \widehat{r}_k\right)\left(m_t - \widehat{m}_t^*\right)\left(\widehat{r}_t - r_t\right) + \left(\prod_{k=1}^{t-1} \widehat{r}_k\right)\left\{\sum_{b_t} \widehat{m}_t^*\left(\widehat{Q}_t + \widehat{\phi}_t - Q_t\right)\right\}\right].
\end{aligned}$$

where the first line follows by definition, the second by iid observations and the definition of $\widehat{m}_0^*$, and the third by Lemma 1. Note that by construction $\widehat{r}_k$ is bounded and by assumption $\widehat{m}_t^*$ is bounded. Therefore, for the second summand in the final display above, Hölder's inequality, Lemmas 7 and 8, the triangle inequality, a Taylor expansion for $\widehat{r}_t - r_t$, and Cauchy-Schwarz yield

$$\left|\sum_{t=1}^{T} \mathbb{E}\left[\left(\prod_{k=1}^{t-1} \widehat{r}_k\right)\left\{\sum_{b_t} \widehat{m}_t^*\left(\widehat{Q}_t + \widehat{\phi}_t - Q_t\right)\right\}\right]\right| \lesssim \sum_{t=1}^{T} \|\widehat{\pi}_t - \pi_t\|^2.$$

Meanwhile, the first summand from the final line in the initial display above can be bounded iteratively, which we consider next.

Arbitrary t:
Beginning with arbitrary $t \in \{1, \ldots T\}$, we have

$$\mathbb{E}\left\{\left(\prod_{k=1}^{t-1} \widehat{r}_k\right)\left(m_t - \widehat{m}_t^*\right)\left(\widehat{r}_t - r_t\right)\right\} = \mathbb{E}\left\{\left(\prod_{k=1}^{t-1} \widehat{r}_k\right)\left(m_t - \widetilde{m}_t^* + \widetilde{m}_t^* - \widehat{m}_t^*\right)\left(\widehat{r}_t - r_t\right)\right\}$$

by adding and subtracting $\widetilde{m}_t^*$. Hölder's inequality, Taylor expansion, and Cauchy-Schwarz yields

$$\left|\mathbb{E}\left\{\left(\prod_{k=1}^{t-1}\widehat{r}_k\right)\left(\widetilde{m}_t^*-\widehat{m}_t^*\right)\left(\widehat{r}_t-r_t\right)\right\}\right|\lesssim\|\widehat{m}_t^*-\widetilde{m}_t^*\|\|\widehat{\pi}_t-\pi_t\|.$$

Meanwhile, for the remaining term,

$$\left|\mathbb{E}\left\{\left(\prod_{k=1}^{t-1}\widehat{r}_k\right)\left(m_t-\widetilde{m}_t^*\right)\left(\widehat{r}_t-r_t\right)\right\}\right|\lesssim\|\widehat{\pi}_t-\pi_t\|\left|\mathbb{E}\left(m_t-\widetilde{m}_t^*\mid A_t,H_t\right)\right|.$$

Because $\widetilde{m}_t^*=\mathbb{E}\{\widehat{P}_{t+1}^*(Z)\mid A_t,H_t\}$ by definition, Lemma 1 dictates that

$$\widetilde{m}_t^*-m_t=\sum_{s=t+1}^{T}\mathbb{E}\left\{\left(\prod_{k=1}^{s-1}\widehat{r}_k\right)\left(m_s-\widehat{m}_s^*\right)\left(\widehat{r}_s-r_s\right)\right\}+\mathbb{E}\left[\left(\prod_{k=1}^{s-1}\widehat{r}_k\right)\left\{\sum_{b_s}\widehat{m}_s^*\left(\widehat{Q}_s+\widehat{\phi}_s-Q_s\right)\right\}\right].$$

Recursion argument:
Because the argument above can be applied to arbitrary t, it can be applied recursively from $t=T$ backwards to $t=1$. The additional doubly robust terms that arise from $m_t^*-m_t$ will be at least as small (asymptotically) as terms that have already appeared in the error. This yields

$$\left|\sum_{t=1}^{T}\mathbb{E}\left[\left(\prod_{k=1}^{t-1}\widehat{r}_k\right)\left(m_t-\widehat{m}_t^*\right)\left(\widehat{r}_t-r_t\right)\right]\right|\lesssim\sum_{t=1}^{T}\|\widehat{m}_t^*-\widetilde{m}_t^*\|\|\widehat{\pi}_t-\pi_t\|.$$

□

D.7 Corollary 2

Proof. This follows by the same argument as for Corollary 1. □

E Proofs for Appendix A

E.1 Proposition 5

Proof. We proceed using the same proof technique as for Proposition 4. Specifically, we'll show that $\mathbb{E}\{\widehat{\varphi}_D(Z)+\widehat{m}_0-m_0\}$ is second order.

The algebra works similarly to the proof for the debiased pseudo-outcomes in Lemma 1. Starting with the summand at the final timepoint, we have

$$\begin{aligned}\mathbb{E}\{\widehat{\varphi}_D(Z)\}&=\mathbb{E}\left[\left(\prod_{s=1}^{T-1}\widehat{r}_s(A_s\mid H_s)\right)\left\{\widehat{Q}_T(1\mid H_T)+\widehat{\phi}_T(1;A_T,H_T)-\widehat{m}_{T-1}(A_{T-1},H_{T-1})\right\}\right]\\&=\mathbb{E}\left[\left(\prod_{s=1}^{T-1}\widehat{r}_s(A_s\mid H_s)\right)\mathbb{E}\left\{\widehat{Q}_T(1\mid H_T)+\widehat{\phi}_T(1;A_T,H_T)-Q_T(1\mid H_T)\mid A_{T-1},H_{T-1}\right\}\right]\\&+\mathbb{E}\left[\left(\prod_{s=1}^{T-1}\widehat{r}_s(A_s\mid H_s)\right)\{Q_T(1\mid H_T)-\widehat{m}_{T-1}(A_{T-1},H_{T-1})\}\right]\end{aligned}$$

$$= \mathbb{E}\left[\left(\prod_{s=1}^{T-1} \widehat{r}_s(A_s \mid H_s)\right) \mathbb{E}\left\{\widehat{Q}_T(1 \mid H_T) + \widehat{\phi}_T(1; A_T, H_T) - Q_T(1 \mid H_T) \mid A_{T-1}, H_{T-1}\right\}\right]$$

$$+ \mathbb{E}\left[\left(\prod_{s=1}^{T-2} \widehat{r}_s(A_s \mid H_s)\right) (\widehat{r}_{T-1} - r_{T-1}) \left\{m_{T-1}(A_{T-1} \mid H_{T-1}) - \widehat{m}_{T-1}(A_{T-1}, H_{T-1})\right\}\right]$$

$$+ \mathbb{E}\left[\left(\prod_{s=1}^{T-2} \widehat{r}_s(A_s \mid H_s)\right) r_{T-1}(A_{T-1} \mid H_{T-1}) \left\{m_{T-1}(A_{T-1} \mid H_{T-1}) - \widehat{m}_{T-1}(A_{T-1}, H_{T-1})\right\}\right].$$

Notice that the first two lines in the final expression are second-order terms. Then, combining the final term with the next summand in the efficient influence function, and omitting arguments, yields

$$\mathbb{E}\left[\left(\prod_{s=1}^{T-2} \widehat{r}_s\right) \left\{r_{T-1}\left(m_{T-1} - \widehat{m}_{T-1}\right) + \sum_{b_{T-1}} \widehat{m}_{T-1}(\widehat{Q}_{T-1} + \widehat{\phi}_{T-1}) - \widehat{m}_{T-2}\right\}\right]$$

$$= \mathbb{E}\left[\left(\prod_{s=1}^{T-2} \widehat{r}_s\right) \left\{\sum_{b_{T-1}} \widehat{m}_{T-1}(\widehat{Q}_{T-1} + \widehat{\phi}_{T-1} - Q_{T-1}) + m_{T-2} - \widehat{m}_{T-2}\right\}\right].$$

Then, we can repeat this algebra $t-2$ more times, and notice that the final $m_0 - \widehat{m}_0$ cancels out so that $\mathbb{E}\{\widehat{\varphi}_D(Z)\} + \widehat{m}_0 - m_0$ is a second-order product of errors, because $\mathbb{E}\{\widehat{Q}_s(b_s \mid H_s) + \widehat{\phi}(b_s; A_s, H_s) - Q_s(b_s \mid H_s) \mid H_s\}$ is a second-order product of errors, by Lemma 8. □

E.2 Lemma 2

Proof. That the pseudo-outcome is unbiased follows by iterated expectations. Meanwhile, one can analyze $\mathbb{E}\{\widehat{P}^*_{t+1}(Z) \mid A_t, H_t\}$ in the same manner as in the proof of Proposition 5, above, with the exception that the product term over $\widehat{r}$ starts at $s = t+1$ rather than $s = 1$. By the final step one has the second-order term in the statement of the result plus $m_t(A_t, H_t)$, which cancels with the $-m_t(A_t, H_t)$ on the left-hand side of (15). □