Longitudinal trimming and smooth trimming with flip and S-flip interventions

Alec McClean* and Iván Díaz

Division of Biostatistics,
New York University Grossman School of Medicine

Abstract

Propensity score trimming addresses positivity violations by excluding individuals with extreme scores. While well-studied with single-timepoint data, standard methods only recommend baseline trimming with longitudinal data. Consequently, trimmed effects remain susceptible to positivity violations at subsequent timepoints. In this paper, we extend trimming to longitudinal data and demonstrate how to trim on non-baseline covariates by leveraging dynamic stochastic interventions. We introduce "flip" interventions, which maintain the treatment status of subjects who would have received the target treatment while flipping others' treatment to the target if they are in the trimmed set. With single-timepoint data, differences in flip effects yield standard trimmed effects. With longitudinal data, they provide interpretable trimming on non-baseline covariates. Crucially, flip interventions are policy-relevant, since they could actually be implemented in practice. We show that other approaches for trimming on non-baseline covariates do not retain this property. We further develop smooth flip ("S-flip") interventions, which incorporate a smooth approximation of the trimming indicator to produce smooth trimmed effects. We derive efficient influence functions for S-flip effects and construct multiply robust-style and sequentially doubly robust-style estimators, which achieve root-n consistency and asymptotic normality under nonparametric conditions.

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

1 Introduction

There is a large and growing literature in causal inference for estimating treatment effects under violations of the positivity assumption. With a binary treatment, the positivity assumption asserts that subjects in every stratum of the covariates have a non-zero probability

*Corresponding author: hadera01@nyu.edu

of receiving treatment and control [Hernán and Robins, 2020]. With longitudinal binary treatments, it asserts that every subject has a non-zero probability of receiving each treatment regime under consideration. Since the number of possible treatment regimes increases exponentially with the number of timepoints, longitudinal data can be very susceptible to positivity violations if one wants to examine effect under all possible treatment regimes. When positivity is fully violated, such that the probability of receiving a particular treatment regime is zero, then the relevant causal effect is not identifiable or estimable from observed data and point estimates for it can be biased. Meanwhile, even if positivity holds estimation can still be challenging — when propensity scores (the probability of receiving a particular treatment) are close to zero (often referred to as "practical" positivity violations) this can inflate variance estimates. These two phenomena ultimately hamper the ability to draw scientifically meaningful conclusions from data [Kang and Schafer, 2007, Khan and Tamer, 2010, Moore et al., 2012, Petersen et al., 2012, Westreich and Cole, 2010].

The causal inference literature has developed two connected approaches for defining effects that are robust to positivity violations: trimming and dynamic stochastic interventions.[1] The first approach, trimming, removes subjects with extreme propensity scores from the analysis [Crump et al., 2009, Frölich, 2004, Smith and Todd, 2005]. Trimming has been extensively studied with single-timepoint data, with recent research developing trimming with continuous treatment and trimming based on conditional variances of potential outcomes [Branson et al., 2023, Khan and Ugander, 2022]. Moreover, Yang and Ding [2018] introduced *smooth* trimmed effects, which use a smooth approximation of the trimming indicator and therefore are pathwise differentiable and, importantly, can be estimated at $\sqrt{n}$-rates under nonparametric assumptions. However, trimming methods have seen only limited extension to longitudinal settings, with existing approaches recommending trimming using baseline covariates [Jensen et al., 2024, Petersen et al., 2012]. As a result, these methods remain susceptible to positivity violations that occur on non-baseline covariates.

By contrast, dynamic stochastic interventions and modified treatment policies (MTPs) offer an alternative approach to addressing positivity violations. These interventions shift the probability of treatment receipt and adapt to positivity violations at each timepoint, ensuring that subjects with zero probability of receiving a treatment are not intervened upon to receive that treatment. Robins et al. [2004] and Stock [1989] pioneered dynamic stochastic interventions dependent on natural treatment values, later examined by Young et al. [2014]. In other early related work van der Laan and Petersen [2007] introduced "realistic" interventions which adapt to positivity violations. Later, Haneuse and Rotnitzky [2013] and Díaz and van der Laan [2012] introduced MTPs and stochastic variants of them

[1] Another approach is truncation, which truncates estimated propensity scores within an interval. It is related to trimming, but different in the sense that the estimand itself is defined by modifying the data-dependent estimated propensity score. For this reason, we leave an examination of truncation to future work.

for single timepoint data, while Taubman et al. [2009] developed threshold interventions. Recent contributions include the incremental propensity score intervention (IPSI) Kennedy [2019], the multiplicative risk ratio intervention Wen et al. [2023], the general framework for dynamic stochastic interventions with unordered multi-valued treatments in McClean et al. [2024b], the extension of IPSIs to continuous treatment in Schindl et al. [2024], longitudinal MTPs [Díaz et al., 2023], and super-optimal MTPs [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

While methods for trimming and dynamic stochastic interventions have developed largely in parallel, in Section 2 we show that there is a straightforward connection between them in single-timepoint data. We demonstrate that trimmed effects can be defined as differences in "flip" interventions, which intervene on subjects that would not have taken the target treatment but are in the trimmed set, flipping them to the target treatment. Although straightforward, this connection between trimmed effects and dynamic interventions has not, to our knowledge, been systematically investigated previously. The link proves particularly valuable because it extends to smooth trimmed effects and longitudinal data. Indeed, in Section 2 we also define smooth flip ("S-flip") interventions, whose differences yield smooth trimmed effects. S-flip interventions intervene on subjects that would not have taken the target treatment and flip them to the target treatment with probability S, where S is the smooth approximation of the trimming indicator.

Building on this insight from single-timepoint data, Section 3 introduces our notation for longitudinal data, while Section 4 provides our main extension of flip and S-flip interventions to longitudinal data. These interventions trim subjects using non-baseline covariates and flip them toward a target treatment regimen. They are particularly useful because they can trim on non-baseline covariates and remain identifiable under positivity violations at all timepoints. They have an additional property: they are "single-world"', in the sense that they rely solely on data observable under the intervention itself. As a result, they are policy-relevant, because they correspond to a policy that could be implemented in practice.

However, our analysis highlights an interesting consequence of the single-world construction of flip interventions: contrasts between two flip effects targeting two regimes can reflect differences in sequential trimming sets as well as capturing mechanistic differences in potential outcomes under the two target regimes. We illustrate this with a simple example in Section 4.2. Then, we consider an alternative trimmed effect that isolates the mechanistic difference in potential outcomes under two regimes and conducts trimming. Crucially, we establish that this effect has major drawbacks: it corresponds to interventions that depend on counterfactual and future propensity scores, and therefore lacks practicality and interpretability.

Section 5 then develops multiply robust-style and sequentially doubly robust-style estimators for the single-world S-flip effects, leveraging their efficient influence functions. These estimators achieve parametric convergence rates under nonparametric conditions, allowing asymptotically valid inference via Gaussian limiting distributions. We establish two new results. First, we show how the bias of the multiply robust-style estimator can be bounded by the minimum of two error decompositions, tightening prior results on the structure of these errors. In addition, our sequentially doubly robust-style guarantee is the first such result for a dynamic stochastic intervention that depends on the unknown propensity score. In ongoing work, we will illustrate our methods using longitudinal data on corticosteroid use and COVID-19 mortality.

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 the notation $a \lesssim b$ to mean $a \leq Cb$ for some constant C, $\rightsquigarrow$ 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 trimming and flip interventions

Before generalizing to longitudinal trimming, we build intuition in the single-timepoint case. 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. Moreover, $Y(D)$ is the potential outcome under treatment decision D, which is a random variable that can depend on the natural value of treatment A and covariates X. Finally, we let $\pi(X) = \mathbb{P}(A = 1 \mid X)$ denote the propensity score.

With one timepoint, a commonly used trimmed average treatment effect (ATE) is $\mathbb{E}\Big[\big\{Y(1) - Y(0)\big\}\mathbb{1}\big\{\varepsilon < \pi(X) < 1 - \varepsilon\big\}\Big]$ where $\varepsilon \geq 0$ is a user-specified threshold, which restricts attention to subjects with propensity scores between ε and $1 - \varepsilon$. We focus on *unconditional* trimmed effects of the form $\mathbb{E}\{f(Z)\mathbb{1}(\mathcal{E})\}$, where $\mathcal{E}$ is an event, because they can be related to differences in interventions across the whole population, as we show next. Moreover, in the single-timepoint case, it is also straightforward to relate unconditional trimmed effect to conditional trimmed effects because $\mathbb{E}\big\{f(Z) \mid \mathcal{E}\big\} = \frac{\mathbb{E}\{f(Z)\mathbb{1}(\mathcal{E})\}}{\mathbb{P}(\mathcal{E})}$ and the methods we develop to identify and estimate the numerator $\mathbb{E}\{f(Z)\mathbb{1}(\mathcal{E})\}$ will also apply to the denominator $\mathbb{P}(\mathcal{E}) = \mathbb{E}\{\mathbb{1}(\mathcal{E})\}$.

Our first result shows that the unconditional trimmed ATE can be defined as the difference in potential outcomes under two flip interventions.

Proposition 1. *Let* $D(1) = A + (1 - A)\mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\}$ *and* $D(0) = A\big[1 - \mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\}\big]$ *denote treatment decisions targeting treatments* $a = 1$ *and* $a = 0$*, respectively. If* $Y(a) \perp\!\!\!\perp A \mid X$ *for* $a \in \{0, 1\}$*, then*

$$\mathbb{E}\Big[Y\{D(1)\} - Y\{D(0)\}\Big] = \mathbb{E}\big[\{Y(1) - Y(0)\}\mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\}\big]. \tag{1}$$

All proofs are delayed to the appendix. Notice that $D(a)$ can also be defined as $D(a) = \mathbb{1}(A = a)A + \mathbb{1}(A \neq a)\Big((1 - A)\mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\} + A\big[1 - \mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\}\big]\Big)$. In other words, $D(a)$ does not intervene if A already equals the target treatment a; otherwise, it flips treatment to $a = 1 - A$ if the subject is in the trimmed set. Therefore, Proposition 1 shows that the trimmed ATE is a difference in interpretable flip interventions.

Smooth trimming. The parallels between trimmed effects and dynamic stochastic interventions are particularly valuable for constructing smooth trimmed effects, which employ a smooth approximation of the trimming indicator function. Smooth trimmed effects have gained popularity because they can be estimated at $\sqrt{n}$-convergence rates under nonparametric assumptions – a property that may not hold for non-smooth effects, as the non-smoothness of the trimming indicator creates a first-order dependence on the propensity score estimator. We define the smooth trimmed ATE as $\mathbb{E}\big[\{Y(1) - Y(0)\}S(X)\big]$, where $S(X)$ is a smooth approximation of $\mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\}$. The next result demonstrates this smooth trimmed ATE can be reformulated as a difference between S-flip effects.

Proposition 2. *Let* $D'(1) = A + (1 - A)\mathbb{1}\{V \leq S(X)\}$ *and* $D'(0) = A\mathbb{1}\{V > S(X)\}$ *where* $V \sim \text{Unif}(0, 1)$ *and* $V \perp\!\!\!\perp (X, A, Y(0), Y(1))$*. If* $Y(a) \perp\!\!\!\perp A \mid X$ *for* $a \in \{0, 1\}$*, then*

$$\mathbb{E}\big[Y\{D'(1)\} - Y\{D'(0)\}\big] = \mathbb{E}\big[\{Y(1) - Y(0)\}S(X)\big].$$

In this case, $D'(a)$ is also defined as $D'(a) = \mathbb{1}(A = a)A + \mathbb{1}(A \neq a)\Big[(1 - A)\mathbb{1}\{V \leq S(X)\} + A\mathbb{1}\{V > S(X)\}\Big]$. In words, this treatment decision says *if* $A = a$*, do not intervene; otherwise, flip treatment with probability* $S(X)$. Importantly, this framework generalizes to longitudinal data, as we show next.

Remark 1. In Proposition 2, 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].

3 Setup and background for longitudinal trimming

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 $\{Z_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P} \in \mathcal{P}$. Each observation will consist 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 $Y\left(\overline{D}_T\right) = f_Y\big\{D_T, H_T(\overline{D}_{T-1}), U_Y\big\}$. Causal effects will be defined in terms of the distribution of this counterfactual random variable.

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). $U_{A,t} \perp\!\!\!\perp \left\{\underline{U}_{X,t+1}, U_Y\right\} \mid H_t$ for all $t \leq T$.

Assumption 2 (Strong sequential randomization). $U_{A,t} \perp\!\!\!\perp \left\{\underline{U}_{X,t+1}, \underline{U}_{A,t+1}, U_Y\right\} \mid H_t$ 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 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.

4 Longitudinal trimming and smooth trimming with flip and S-flip interventions

In this section, we extend flip and S-flip interventions from Section 2 to longitudinal data. We define longitudinal flip and S-flip interventions and identify the resulting effects, which perform (smooth) trimming on non-baseline covariates. These effects are identifiable under arbitrary positivity violations; moreover, they are "single-world," in the sense that they only rely on information that would be observed under the series of interventions. As a result, they retain policy-relevance, as they could be implemented in practice. However, we also illustrate a subtle consequence of this construction: differences in single-world flip effects can be driven not only by differences in potential outcomes under the two targeted regimes, but also by differences in non-baseline trimming. Therefore, we also introduce an alternative trimmed effect that isolates only the contrast in potential outcomes. We show that this alternative effect corresponds to interventions that rely on trimming via counterfactual, or "cross-world," propensity scores and future propensity scores. Thus, despite their theoretical appeal in isolating contrasts between two potential outcomes, we show these cross-world future-dependent trimmed effects have serious practical and interpretational limitations.

4.1 Single-world flip and S-flip interventions

We now present our primary proposal for longitudinal trimming and smooth trimming: single-world flip and S-flip interventions. These extend the interventions in Section 2 to longitudinal data. We begin by defining them and then establish conditions under which the resulting effects remain identifiable without standard positivity assumptions.

Definition 1 (Single-world flip interventions)**.** 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

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

where

$$I_t^c(a_t; h_t) = \mathbb{1}\big[\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1}) = h_t\} > \varepsilon\big].$$

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 a_t if the propensity score they *would have had* for treatment a_t exceeds ε.

For brevity, we suppress the explicit dependence of D_t on the target treatment a_t, prior interventions $\overline{D}_{t-1}$, and the natural values of treatment and history $\{A_t(\overline{D}_{t-1}), H_t(\overline{D}_{t-1})\}$, although one could write $D_t = d_t\{a_t; \overline{D}_{t-1}, A_t(\overline{D}_{t-1}), H_t(\overline{D}_{t-1}\}$ where d_t is the function above. A further important nuance is that the trimming indicator $I_t^c(\cdot)$ relies on propensity scores determined by both the natural covariate history and the natural treatment value. Consequently, the intervention is "single-world," meaning the decision at time t depends only on information observable under interventions performed up to that point ($\overline{D}_{t-1}$). The superscript "c" emphasizes the "counterfactual" nature of $I_t^c(\cdot)$.

As in the single-timepoint setting, flip interventions involve an indicator function that is non-smooth, complicating nonparametric estimation. Consequently, an estimator of a flip effect typically has a first-order dependence on the estimated propensity score. We therefore introduce *S-flip interventions* to generalize *smooth trimming* to longitudinal data, which can yield $\sqrt{n}$-estimable effects under nonparametric conditions.

Definition 2 (S-flip intervention)**.** Let $\overline{a}_T = \{a_1, \ldots, a_T\} \in \{0,1\}^T$ be the target regime. A *smooth flip (S-flip) intervention* at time t targeting a_t is

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

where $V_1, \ldots, V_T$ are i.i.d. Unif$(0,1)$ random variables with $V_t \perp\!\!\!\perp Z$ for all $t \in \{1, \ldots, T\}$, and

$$S_t^c(a_t; h_t) = s\Big[\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\}; \varepsilon_t, k_t\Big]$$

where $s(\cdot; \varepsilon_t, k_t)$ is a smooth approximation to $\mathbb{1}(x > \varepsilon)$ and k_t is a smoothing parameter.

In words, at time t:

- if the natural value of treatment is already a_t, the S-flip intervention does nothing;
- otherwise, it "flips" the subject to the target treatment a_t with probability $S_t^c\{a_t; H_t(\overline{D}_{t-1})\}$, where S_t is a smooth approximation of the indicator function that the propensity score they *would have had* for treatment a_t exceeds ε.

For now, we leave the specific smoothing function unspecified. The following result establishes conditions under which flip and S-flip effects remain identifiable, even under arbitrary positivity violations. Afterward, we introduce two simple smoothing functions that enable identification under these conditions.

Theorem 1 (Identification of flip and S-flip effects)**.** *Let $\overline{D}_T = \{D_1, D_2, \ldots, D_T\}$ denote flip interventions as in Definition 1 targeting treatment regime $\overline{a}_T$. Suppose the NPSEM and Assumption 2 hold. Then,*

$$\mathbb{E}\{Y(\overline{D}_T)\} = \sum_{\overline{b}_T \in \{0,1\}^T} \mathbb{E}\left\{\mathbb{E}\big(Y \mid \overline{A}_T = \overline{b}_T,\ \overline{X}_T\big) \prod_{t=1}^{T} Q_t\big(b_t \mid \overline{b}_{t-1}, \overline{X}_t\big)\right\}, \tag{2}$$

$$= \mathbb{E}\left[Y \prod_{t=1}^{T} \frac{Q_t(A_t \mid H_t)}{\mathbb{P}(A_t \mid H_t)}\right], \tag{3}$$

where

$$Q_t(b_t \mid h_t) \ = \ \begin{cases} \mathbb{P}(A_t = b_t \mid h_t)\,\big\{1 - I_t(a_t; h_t)\big\} + I_t(a_t; h_t), & \text{if } b_t = a_t, \\ \mathbb{P}(A_t = b_t \mid h_t)\,\big\{1 - I_t(a_t; h_t)\big\}, & \text{otherwise,} \end{cases}$$

and

$$I_t(a_t; h_t) = \mathbb{1}\{\mathbb{P}(A_t = a_t \mid H_t = h_t) > \varepsilon\}.$$

With the same assumptions, let $\overline{D}_T$ be S-flip interventions as in Definition 2, satisfying $\mathbb{P}\{A_t(\overline{D}_{t-1}) = a_t \mid H_t(\overline{D}_{t-1})\} = 0 \implies S_t^c\{a_t; H_t(\overline{D}_{t-1})\} = 0$. Then (2) and (3) also hold with

$$Q_t(b_t \mid h_t) \ = \ \begin{cases} \mathbb{P}(A_t = b_t \mid h_t)\,\big\{1 - S_t(a_t; h_t)\big\} + S_t(a_t; h_t), & \text{if } b_t = a_t, \\ \mathbb{P}(A_t = b_t \mid h_t)\,\big\{1 - S_t(a_t; h_t)\big\}, & \text{otherwise,} \end{cases}$$

where

$$S_t(a_t; h_t) = s\{\mathbb{P}(A_t = a_t \mid H_t = h_t); \varepsilon_t, k_t\}.$$

Theorem 1 establishes that mean potential outcomes under flip and S-flip interventions are identifiable under only strong sequential randomization and consistency. Equation (2) provides the g-formula identification [Robins, 1986] while equation (3) provides the inverse weighting identification. For flip interventions, the result requires that the trimming indicator is zero when the propensity score for the target treatment is zero; similarly, for S-flip interventions, the smooth trimming indicator must be zero. These constraints prevent positivity violations that would render the causal effect undefined and can be enforced through appropriate construction of the trimming indicators. For instance, when $\varepsilon = 0$, McClean et al. [2024b] proposed a simple smooth trimming indicator: $s(x; 0, k) = 1 - \exp(-kx)$ for $k > 0$. For trimming with $\varepsilon > 0$, one might consider $s(x; \varepsilon, k) = \frac{x}{x+\exp\{-k(x-\varepsilon)\}}$ for $k > 0$.

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

1. **Robustness to positivity violations.** S-flip effects retain robustness to positivity violations, making them a valuable alternative to incremental propensity score interventions (IPSIs) [Bonvini et al., 2023, Kennedy, 2019]. While both IPSIs and S-flip interventions are time-varying dynamic stochastic interventions that remain identifiable under positivity violations, S-flip interventions are distinct in their ability 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 and S-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 and S-flip interventions depend on the natural treatment value to retain an intuitive interpretation and intervene on as few subjects as possible. 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, for flip interventions, one could instead define

$$D_t = \mathbb{1}\Big(V_t \leq \mathbb{1}(a_t = 1) I_t^c\{a_t; H_t(\overline{D}_{t-1})\}$$

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

where $V_t \sim \text{Unif}(0,1)$. Similarly, for S-flip interventions, one could use

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

These redefined interventions satisfy the identification result of Theorem 1 under standard sequential randomization and do not suffer from the practical concerns discussed in the previous point.

4. **Connections to maximally coupled policies.** Flip and S-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 Contrasts of flip effects

In this section, we investigate the properties of contrasts of single-world flip effects. First, we note that they satisfy a minimal property: if the treatment has no effect on the outcome then the contrast of single-world flip effects equals zero.

Proposition 3. *Let $\overline{D}_T$ and $\overline{D}'_T$ denote two flip interventions. If $Y(\overline{b}_T) = Y(\overline{b}'_T)$ for all $\overline{b}_T, \overline{b}'_T \in \{0,1\}^T$, then $\mathbb{E}\{Y(\overline{D}_T) - Y(\overline{D}'_T)\} = 0$.*

However, there is also a subtle consequence of the *single-world* nature of flip interventions: differences between flip effects under two target regimes may reflect both mechanistic differences in potential outcomes and differences in sequential trimming. In essence, 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 $Y(\overline{D}_T)$. The following example illustrates this point in more detail.

Example 1. Suppose $T = 2$, $\overline{D}_2$ is a single-world flip intervention targeting always-treated, and $\overline{D}'_2$ is a single-world flip interventions targeting never-treated. Suppose there are no positivity violations at baseline. Then, D_1 pushes *all subjects* into treatment (i.e., $\mathbb{P}\{\mathbb{P}(D_1 = 1 \mid X_1) = 1\} = 1$) and D'_1 pushes *all subjects* into control (i.e., $\mathbb{P}\{\mathbb{P}(D'_1 = 1 \mid X_1) = 0\} = 1$). Hence, by iterated expectations,

$$\mathbb{E}\{Y(\overline{D}_2) - Y(\overline{D}'_2)\} = \sum_{b_2} \mathbb{E}\left[Y(b_2, 1)\mathbb{P}\{D_2 = b_2 \mid H_2(1)\}\right] - \sum_{b_2} \mathbb{E}\left[Y(b_2, 0)\mathbb{P}\{D'_2 = b_2 \mid H_2(0)\}\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[Y(1,1)\mathbb{P}\{D_2 = 1 \mid H_2(1)\} - Y(0,0)\mathbb{P}\{D_2' = 0 \mid H_2(0)\}\right] \\
&+ \mathbb{E}\left[Y(0,1)\mathbb{P}\{D_2 = 0 \mid H_2(1)\} - Y(1,0)\mathbb{P}\{D_2' = 1 \mid H_2(0)\}\right].
\end{aligned}
$$

There are two properties worth emphasizing:

1. Suppose trimming is different depending whether one is treated or not in the first timepoint, i.e., $\mathbb{P}\{D_2 = 1 \mid H_2(1)\}$ and $\mathbb{P}\{D_2' = 0 \mid H_2(0)\}$ differ. Then, $\mathbb{E}\left[Y(1,1)\mathbb{P}\{D_2 = 1 \mid H_2(1)\} - Y(0,0)\mathbb{P}\{D_2' = 0 \mid H_2(0)\}\right]$ does **not** simplify to reflect only a weighted difference of $Y(1,1) - Y(0,0)$; it also reflects differences in second-timepoint trimming.

2. Suppose there are any positivity violations at the second timepoint, i.e., $\mathbb{P}\big[\mathbb{P}\{D_2 = 0 \mid H_1(1)\} > 0\big] > 0$ or $\mathbb{P}\big[\mathbb{P}\{D_2' = 1 \mid H_2(0)\} > 0\big] > 0$. Then, the difference $\mathbb{E}\{Y(\overline{D}_2) - Y(\overline{D}_2')\}$ will depend on $Y(0,1)$ and $Y(1,0)$ in addition to the difference $Y(1,1) - Y(0,0)$.

This example highlights two ways that contrasts of single-world flip interventions may not purely reflect mechanistic differences between potential outcomes under the target treatment regimes. The first property arises because the non-baseline interventions are trimming using different propensity scores. Meanwhile, the second property arises because a shrinking trimmed set means the ultimate effects consider mean potential outcomes under non-target regimes; e.g., $\mathbb{E}\{Y(\overline{D}_2)\}$ incorporates $Y(0,1)$ due to positivity violations at the second timepoint. These are not necessarily issues, since each property arises because the single-world flip interventions are deliberately constructed to be policy-relevant and implementable in practice. Indeed, both properties can be removed, but require constructing interventions using ***cross-world propensity scores*** (for the first issue) and ***future propensity scores*** (for the second issue), as we show next.

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

$$
\psi(\overline{a}_T, \overline{a}_T') = \mathbb{E}\Big[\{Y(\overline{a}_T) - Y(\overline{a}_T')\} \prod_{t=1}^{T} \mathbb{1}\Big\{\mathbb{P}\big(A_t(\overline{a}_{t-1}) = a_t \mid H_t(\overline{a}_{t-1})\big) > 0,\ \mathbb{P}\big(A_t(\overline{a}_{t-1}') = a_t' \mid H_t(\overline{a}_{t-1}')\big) > 0\Big\}\Big]. \quad (4)
$$

This effect isolates the difference $Y(\overline{a}_T) - Y(\overline{a}_T')$ among subjects who would have non-zero probability of receiving both regimes. As a result, it preserves the stochastic ordering of $Y(\overline{a}_T)$ and $Y(\overline{a}_T')$. However, it has two major limitations. First, it is "cross-world," meaning the trimming function depends on counterfactual covariates under unobservable treatment regimes. Consequently, it 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, it corresponds to a contrast under *future-dependent* interventions, as clarified by Proposition 4, next.

Proposition 4. *Let* $\Pi_T := \prod_{t=1}^{T} \mathbb{1}\Big\{\mathbb{P}\big(A_t(\overline{a}_{t-1}) = a_t \mid H_t(\overline{a}_{t-1})\big) > 0,\ \mathbb{P}\big(A_t(\overline{a}'_{t-1}) = a'_t \mid H_t(\overline{a}'_{t-1})\big) > 0\Big\}$*. Then, the treatment decisions* $\overline{D}_T = \mathbb{1}(\overline{A}_T = \overline{a}_T)\overline{A}_T + \mathbb{1}(\overline{A}_T \neq \overline{a}_T)\Big\{\overline{a}_T\Pi_T + \overline{A}_T\big(1-\Pi_T\big)\Big\}$ *and* $\overline{D}'_T = \mathbb{1}(\overline{A}_T = \overline{a}'_T)\overline{A}_T + \mathbb{1}(\overline{A}_T \neq \overline{a}'_T)\Big\{\overline{a}'_T\Pi_T + \overline{A}_T\big(1-\Pi_T\big)\Big\}$ *yield* $\mathbb{E}\Big\{Y(\overline{D}_T) - Y(\overline{D}'_T)\Big\} = \psi(\overline{a}_T, \overline{a}'_T)$.

Proposition 4 shows that the trimmed effect in (4) is defined as differences in mean potential outcomes under two treatment decisions. These treatment decisions correspond to *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 at 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 nature of the intervention may hamper the interpretability of the effects.

One may wonder whether $\overline{D}_T$ and $\overline{D}'_T$ can be constructed that do not use future information to inform earlier interventions. The next result answers this in the negative, and shows that, when additional trimming happens after the target regimes diverge, then one must use future information when constructing treatment rules $\overline{D}_T$ and $\overline{D}'_T$ to define the trimmed effect in (4).

Theorem 2 (Impossibility)**.** *Let* $\Pi_k := \prod_{t=1}^{k} \mathbb{1}\Big\{\mathbb{P}\big(A_t(\overline{a}_{t-1}) = a_t \mid H_t(\overline{a}_{t-1})\big) > 0,\ \mathbb{P}\big(A_t(\overline{a}'_{t-1}) = a'_t \mid H_t(\overline{a}'_{t-1})\big) > 0\Big\}$*. Suppose*

1. *the target regimes diverge before the final timepoint, i.e.,* $\overline{a}_t \neq \overline{a}'_t$ *for* $t < T$*, and*
2. *there is additional trimming after the target regimes diverge, i.e., there exists* $s > t$ *such that* $\mathbb{P}(\Pi_t) - \mathbb{P}(\Pi_s) > 0$*.*

Then, one cannot construct interventions $\overline{D}_T$ *and* $\overline{D}'_T$ *which only depend on current and past treatment and covariate information yielding* $\mathbb{E}\{Y(\overline{D}_T) - Y(\overline{D}'_T)\} = \psi(\overline{a}_T, \overline{a}'_T)$*. In other words, the interventions* ***must*** *use future information.*

5 Estimation and inference

In this section, we outline methods for estimating single-world S-flip effects as in Definition 2. Throughout, we have assumed that the trimming function or smooth trimming function were 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 the trimming threshold or smooth trimming parameter data-adaptively — then estimation and inference are more complex; see Khan and Ugander [2022] for a review.

Even in our simpler setting where the trimming function is fixed, conducting inference presents different challenges depending on whether we use smooth or non-smooth trimmed effects. For non-smooth trimmed effects, the lack of pathwise differentiability due to the non-smoothness of the trimming indicator 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 trimmed effects, 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 S-flip effect and then we will use it to construct a multiply robust and sequentially doubly robust estimators.

5.1 Notation

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

$$r_t(b_t \mid h_t) = \frac{Q_t(b_t \mid h_t)}{\mathbb{P}(A_t = b_t \mid h_t)} \tag{5}$$

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

$$m_t(b_t, h_t) = \mathbb{E}\left\{ \sum_{b_{t+1}} m_{t+1}(b_{t+1}, H_{t+1}) Q_{t+1}(b_{t+1} \mid H_{t+1}) \mid A_t = b_t, H_t = h_t \right\} \tag{6}$$

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 the S-flip effect.

Proposition 5. *Let ψ denote an identified S-flip effect from Theorem 1 where $s_t(\cdot;\varepsilon,k)$ is twice differentiable with non-zero and bounded derivatives. Moreover, let*

$$\phi_t(b_t;A_t,H_t) = \Big\{\mathbb{1}(A_t=a_t) - \mathbb{P}(A_t=a_t \mid H_t)\Big\} s_t'\{\mathbb{P}(A_t=a_t \mid H_t);\varepsilon_t,k_t\}\Big\{\mathbb{1}(b_t=a_t) - \mathbb{P}(A_t=b_t \mid H_t)\Big\} + \Big\{\mathbb{1}(A_t=b_t) - \mathbb{P}(A_t=b_t \mid H_t)\Big\}\Big[1 - s_t\{\mathbb{P}(A_t=a_t \mid H_t);\varepsilon_t,k_t\}\Big]$$

where $s_t'(x;\varepsilon,k) = \frac{\partial}{\partial x} s_t(x;\varepsilon,k)$. Further suppose that the outcome Y has bounded variance and the S-flip 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

$$\varphi(Z) = \varphi_m(Z) + \varphi_Q(Z) \text{ where}$$

$$\varphi_m(Z) = \sum_{t=0}^{T}\left\{\prod_{s=0}^{t} r_s(A_s \mid H_s)\right\}\left\{\sum_{b_{t+1}} m_{t+1}(b_{t+1},H_{t+1})Q_{t+1}(b_{t+1} \mid H_{t+1}) - m_t(A_t,H_t)\right\},$$

$$\varphi_Q(Z) = \sum_{t=1}^{T}\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\}\sum_{b_t} m_t(b_t,H_t)\phi_t(b_t;A_t,H_t).$$

The efficient influence function in Proposition 5 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. For the S-flip effect in Proposition 5, $\varphi_Q(Z)$ mirrors the form found in the IPSI [Kennedy, 2019, Theorem 2].

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 smooth trimming indicator. 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 5 naturally 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 $\mathbf{Z}_k$ 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 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 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*

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

where

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

Finally, output the point estimate and variance estimates

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

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 3. *Under the setup of Proposition 5, 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 $\mathbb{P}\big\{\widehat{m}_t(A_t, H_t) \leq C\big\} = \mathbb{P}\big\{m_t(A_t, H_t) \leq C\big\} = 1$ 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 3 provides a crucial bound on the bias of the multiply robust-style estimator. Under the assumptions of Proposition 5, 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:

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 $\{\widehat{Q}_s\}_{s \le t}$.

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 $s \ge t$ 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 interventional propensity scores 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.).

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

Corollary 1 (Multiple robustness and weak convergence)**.** *Under the setup of Theorem 3, 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

$$\sqrt{\frac{n}{\widehat{\sigma}^2}} \big(\widehat{\psi} - \psi\big) \rightsquigarrow N(0,1). \tag{7}$$

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 3 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 2. Corollary 1 can be extended to a range of smoothing parameters $\{k_1, \ldots, k_K\}$, yielding a uniform convergence result as in Kennedy [2019, Theorem 3], under slightly stronger assumptions than those in Corollary 1. This is feasible provided the efficient influence function $\varphi(Z)$ from Proposition 5 is sufficiently smooth in the smoothing parameter. For instance, the smoothing function $s(x) = 1 - \exp(-kx)$ and other related choices satisfy this requirement.

Remark 3. 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 3 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.

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 5, 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,

$$\mathbb{E}\left\{P_{t+1}^*(Z) \mid A_t, H_t\right\} = m_t(A_t, H_t). \tag{8}$$

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,*

$$\mathbb{E}\left\{\widehat{P}_{t+1}^*(Z) - m_t(A_t, H_t) \mid A_t, H_t\right\} = \sum_{s=t+1}^{T} \mathbb{E}\left[\left\{\prod_{k=t+1}^{s-1} \widehat{r}_k(A_k \mid H_k)\right\}\left\{m_s(A_s, H_s) - \widehat{m}_s^*(A_s, H_s)\right\}\left\{\widehat{r}_s(A_s \mid H_s) - r_s(A_s \mid H_s)\right\} \mid A_t, H_t\right] + \sum_{s=t+1}^{T} \mathbb{E}\left[\left\{\prod_{k=t+1}^{s-1} \widehat{r}_k(A_k \mid H_k)\right\}\sum_{b_s} \widehat{m}_s^*(A_s, H_s)\left\{\widehat{Q}_s(b_s \mid H_s) + \widehat{\phi}_s(b_s; A_s, H_s) - Q_s(b_s \mid H_s)\right\} \mid A_t, H_t\right]. \tag{9}$$

This lemma proposes the debiased pseudo-outcome, P_t^*, then shows that it is indeed unbiased (in (8)) and that its error, if it were estimated, is a product of errors (in (9)). 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 (9) 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

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

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 5 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 4. *Under the setup of Theorem 3, 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 4 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 4, 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

$$\sqrt{\frac{n}{\widehat{\sigma}^2}}\big(\widehat{\psi} - \psi\big) \rightsquigarrow N(0, 1). \tag{10}$$

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 methods for flip and S-flip effects with administrative censoring and censoring by death. We will apply these methods to analyze the effect of corticosteroid use on mortality for patients with moderate to severe COVID-19 using a retrospective cohort of patients at NewYork-Presbyterian Hospital during Spring 2020, at the beginning of the pandemic. Future versions of this manuscript will include this data analysis.

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.

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.

Markus Frölich. Finite-sample properties of propensity-score matching and weighting estimators. *Review of Economics and Statistics*, 86(1):77–90, 2004.

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

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.

Shakeeb Khan and Elie Tamer. Irregular identification, support conditions, and inverse weight estimation. *Econometrica*, 78(6):2021–2042, 2010.

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.

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.

Jeffrey A Smith and Petra E Todd. Does matching overcome lalonde's critique of nonexperimental estimators? *Journal of econometrics*, 125(1-2):305–353, 2005.

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.

Mark J van der Laan and Maya L Petersen. Causal effect models for realistic individualized treatment and intention to treat rules. *The international journal of biostatistics*, 3(1), 2007.

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.

Lan Wen, Julia L Marcus, and Jessica G Young. Intervention treatment distributions that depend on the observed treatment process and model double robustness in causal survival analysis. *Statistical methods in medical research*, 32(3):509–523, 2023.

Daniel Westreich and Stephen R Cole. Invited commentary: positivity in practice. *American journal of epidemiology*, 171(6):674–677, 2010.

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.

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.

Appendix

This appendix contains the following sections:

Appendix A provides proofs for the results in Section 2.

Appendix B provides proofs for the results in Section 4.

Appendix C provides proofs for the results in Section 5.

A Proofs for Section 2: Propositions 1 and 2

Proof. For Proposition 1, we have

$$\begin{aligned}
\mathbb{E}\big[Y\{D(1)\} - Y\{D(0)\}\big] &= \mathbb{E}\Big(\mathbb{E}\big[Y\{D(1)\} \mid X\big]\Big) - \mathbb{E}\Big(\mathbb{E}\big[Y\{D(0)\} \mid X\big]\Big) \\
&= \mathbb{E}\Big\{\mathbb{E}\Big(\mathbb{E}\big[Y\{D(1)\} \mid D(1), X\big] \mid X\Big)\Big\} - \mathbb{E}\Big\{\mathbb{E}\Big(\mathbb{E}\big[Y\{D(0)\} \mid D(0), X\big] \mid X\Big)\Big\} \\
&= \mathbb{E}\Big[\mathbb{E}\big\{Y(1) \mid D(1) = 1, X\big\}\mathbb{P}\{D(1) = 1 \mid X\} \\
&\qquad + \mathbb{E}\big\{Y(0) \mid D(1) = 0, X\big\}\mathbb{P}\{D(1) = 0 \mid X\}\Big] \\
&\quad - \mathbb{E}\Big[\mathbb{E}\big\{Y(1) \mid D(0) = 1, X\big\}\mathbb{P}\{D(0) = 1 \mid X\} \\
&\qquad - \mathbb{E}\big\{Y(0) \mid D(0) = 0, X\big\}\mathbb{P}\{D(0) = 0 \mid X\}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\{Y(1) \mid X\}\mathbb{P}\{D(1) = 1 \mid X\} + \mathbb{E}\{Y(0) \mid X\}\mathbb{P}\{D(1) = 0 \mid X\}\Big] \\
&\qquad - \mathbb{E}\Big[\mathbb{E}\{Y(1) \mid X\}\mathbb{P}\{D(0) = 1 \mid X\} + \mathbb{E}\{Y(0) \mid X\}\mathbb{P}\{D(0) = 0 \mid X\}\Big] \\
&= \mathbb{E}\left(\mathbb{E}\{Y(1) \mid X\}\Big[\mathbb{P}\{D(1) = 1 \mid X\} - \mathbb{P}\{D(0) = 1 \mid X\}\Big]\right) \\
&\qquad + \mathbb{E}\left(\mathbb{E}\{Y(0) \mid X\}\Big[\mathbb{P}\{D(1) = 0 \mid X\} - \mathbb{P}\{D(0) = 0 \mid X\}\Big]\right)
\end{aligned}$$

where the first line follows by iterated expectations on X, the second by iterated expectations on X and $D(1)$ and X and $D(0)$, the third by the definition of expectation, and the fourth by exchangeability and the definitions of $D(1)$ and $D(0)$. Let $\mathbb{1}(X) \equiv \mathbb{1}\{\varepsilon < \pi(X) < 1 - \varepsilon\}$. Then, by the law of total probability over A,

$$\begin{aligned}
&\mathbb{P}\{D(1) = 1 \mid X\} = \pi(X) + \mathbb{1}(X)\{1 - \pi(X)\} = \mathbb{1}(X) + \pi(X)\{1 - \mathbb{1}(X)\}, \\
&\mathbb{P}\{D(1) = 0 \mid X\} = \{1 - \mathbb{1}(X)\}\{1 - \pi(X)\} = 1 - \mathbb{1}(X) - \pi(X)\{1 - \mathbb{1}(X)\}, \\
&\mathbb{P}\{D(0) = 1 \mid X\} = \pi(X)\big\{1 - \mathbb{1}(X)\big\}, \text{ and} \\
&\mathbb{P}\{D(0) = 0 \mid X\} = 1 - \pi(X)\big\{1 - \mathbb{1}(X)\big\}.
\end{aligned}$$

Therefore,

$$\mathbb{P}\{D(1) = 1 \mid X\} - \mathbb{P}\{D(0) = 1 \mid X\} = \mathbb{1}(X)$$

and

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

Plugging these back into the display above yields

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

The result follows by iterated expectations on X.

The same argument follows for Proposition 2, but noting that

$$\begin{aligned}
&\mathbb{P}\{D(1)=1 \mid X\}=\pi(X)+\{1-\pi(X)\}S(X),\\
&\mathbb{P}\{D(1)=0 \mid X\}=\{1-\pi(X)\}\{1-S(X)\},\\
&\mathbb{P}\{D(0)=1 \mid X\}=\pi(X)\{1-S(X)\}, \text{ and}\\
&\mathbb{P}\{D(0)=0 \mid X\}=1-\pi(X)\{1-S(X)\}
\end{aligned}$$

and canceling terms. □

B Proofs for Section 4

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, let

- $\{D_1, D_2(D_1), \ldots, D_T(\overline{D}_{T-1})\}$ 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 $V_1, \ldots, V_T$ are mutually independent and $V_t \perp\!\!\!\perp Z$ for all $t \in \{1, \ldots, T\}$,
- $\overline{X}_t(\overline{a}_{t-1})$ denote the natural covariate history under an intervention that sets treatment to $\overline{a}_{t-1}$ up until time $t-1$,
- $H_t(\overline{a}_{t-1}) = \left\{\overline{X}_{t-1}(\overline{a}_{t-1}), \overline{D}_{t-1} = \overline{a}_{t-1}\right\}$ denote the natural covariate history and the intervention treatment history,
- $A_t(\overline{a}_{t-1})$ denote the natural value of treatment after history $H_t(\overline{a}_{t-1})$,
- $D_t(\overline{a}_{t-1}) = f_t\{A_t(\overline{a}_{t-1}), H_t(\overline{a}_{t-1}), V_t\}$, and
- $Y(\overline{a}_t, \underline{D}_{t+1})$ denote the potential outcome under an intervention that sets treatment $\overline{a}_t$ up until time t and assigns treatment according to treatment decisions $\underline{D}_{t+1} = \{D_{t+1}(\overline{a}_t), D_{t+2}(D_{t+1}, \overline{a}_t), \ldots D_T(\overline{a}_t, D_{t+1}, \ldots, D_{T-1})\}$ thereafter.

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

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

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

Lemma 2. *Under Assumption 2,*

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

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 $(\underline{U}_{A,t+1}, \underline{U}_{X,t+1}, U_Y)$. The first result follows by Assumption 2. The second result follows by Assumption 2 and the assumption on $\overline{V}_T$. □

Lemma 3. *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 and S-flip interventions.*

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

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

By construction, $\mathbb{P}(A_t = a_t \mid H_t) = 0 \implies I_t(a_t; H_t) = 0$; 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 $Q_t(1 - a_t \mid H_t) = 0$.

The same argument holds for S-flip interventions as long as $S_t(a_t; H_t) = 0$ whenever $\mathbb{P}(A_t = a_t \mid H_t) = 0$, which holds by construction of the smooth trimming indicator. □

B.1 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 X_1 and D_1, the second by taking the expectation over D_1, the third by Lemma 2 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 3, 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

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

$$= \mathbb{E}\Big[\sum_{b_2} \mathbb{E}\big\{Y(b_1, b_2, \underline{D}_3) \mid A_2(b_1) = b_2, H_2(b_1)\big\} Q_2(b_2 \mid b_1, \overline{X}_2) \mid A_1 = b_1, X_1\Big]$$

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

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 2 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 3, 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. □

B.2 Proposition 4

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 a deterministic function conditional on $\overline{A}_T, \Pi_T$.

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

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

The difference in propensity scores is just a deterministic function; indeed, for the trimming function Π_T, by the definition of $\overline{D}_T$ and $\overline{D}_T'$, it satisfies

$$\mathbb{P}(\overline{D}_T = \bar{b}_T \mid \overline{A}_T, \Pi_T) - \mathbb{P}(\overline{D}_T' = \bar{b}_T \mid \overline{A}_T, \Pi_T) = \{\mathbb{1}(\bar{b}_T = \overline{a}_T) - \mathbb{1}(\bar{b}_T = \overline{a}_T')\}\Pi_T.$$

Therefore,

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

The result follows by iterated expectations on $\overline{A}_T, \Pi_T$. □

B.3 Theorem 2

Proof. Suppose towards a contradiction that one could define treatment rules that only depend on current and past information. Then, the treatment rules $\{D_1, \ldots, D_T\}$ would yield interventional propensity scores

$$\{\mathbb{P}\{D_t = a_t \mid H_t(\overline{D}_{t-1})\}\}_{t=1}^{T}$$

for $a_t \in \{0,1\}$. By sequentially applying iterated expectations (but not identifying the g-formula), we have

$$\begin{aligned}
&\mathbb{E}\Big\{Y(\overline{D}_T) - Y(\overline{D}'_T)\Big\} \\
&= \sum_{\bar{b}_T} \mathbb{E}\left[Y(\bar{b}_T)\left\{\prod_{t=1}^{T} \mathbb{P}\left\{D_t = b_t \mid H_t(\bar{b}_{t-1})\right\} - \prod_{t=1}^{T} \mathbb{P}\left\{D'_t = b_t \mid H'_t(\bar{b}_{t-1})\right\}\right\}\right].
\end{aligned}$$

In order to guarantee $\mathbb{E}\big\{Y(\overline{D}_T) - Y(\overline{D}'_T)\big\} = \psi(\overline{a}_T, \overline{a}'_T)$, it must be the case that

1. $\prod_{t=1}^{T} \mathbb{P}\{D_t = a_t \mid \overline{H}_t(\overline{a}_{t-1})\} - \prod_{t=1}^{T} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} = \Pi_T$,
2. $\prod_{t=1}^{T} \mathbb{P}\{D_t = a'_t \mid H_t(\overline{a}'_{t-1})\} - \prod_{t=1}^{T} \mathbb{P}\{D'_t = a'_t \mid H_t(\overline{a}'_{t-1})\} = -\Pi_T$, and
3. $\prod_{t=1}^{T} \mathbb{P}\{D_t = b_t \mid H_t(\bar{b}_{t-1})\} - \prod_{t=1}^{T} \mathbb{P}\{D'_t = b_t \mid H_t(\bar{b}_{t-1})\} = 0$ for $\bar{b}_T \notin \{\overline{a}_T, \overline{a}'_T\}$.

In what follows, let $\overline{c}_T = \{a_1, \ldots, 1-a_s, a_{s+1}, \ldots a_T\}$, where s is the timepoint where additional trimming occurs, so that $\mathbb{P}(\Pi_t) - \mathbb{P}(\Pi_s) > 0$. Note that $\overline{c}_T \notin \{\overline{a}_T, \overline{a}'_T\}$ by the assumption of the theorem that $\overline{a}_t \neq \overline{a}'_t$ for $t < T$. Then,

$$\begin{aligned}
\Pi_T &= \prod_{t=1}^{T} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} - \prod_{t=1}^{T} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&= \prod_{t\neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} - \prod_{t\neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&\qquad - \big[1 - \mathbb{P}\{D_s = a_s \mid H_s(\overline{a}_{s-1})\}\big] \prod_{t\neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&\qquad + \big[1 - \mathbb{P}\{D'_s = a_s \mid H_s(\overline{a}_{s-1})\}\big] \prod_{t\neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&= \prod_{t\neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} - \prod_{t\neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&\qquad - \mathbb{P}\{D_s = 1 - a_s \mid H_s(\overline{a}_{s-1})\} \prod_{t\neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&\qquad + \mathbb{P}\{D'_s = 1 - a_s \mid H_s(\overline{a}_{s-1})\} \prod_{t\neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&= \prod_{t\neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} - \prod_{t\neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\} \\
&\qquad - \mathbb{P}\{D_s = c_s \mid H_s(\overline{c}_{s-1})\} \prod_{t\neq s} \mathbb{P}\{D_t = c_t \mid H_t(\overline{c}_{t-1})\} \\
&\qquad + \mathbb{P}\{D'_s = c_s \mid H_s(\overline{c}_{s-1})\} \prod_{t\neq s} \mathbb{P}\{D'_t = c_t \mid H_t(\overline{c}_{t-1})\}
\end{aligned}$$

$$= \prod_{t \neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} - \prod_{t \neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\}$$

$$- \prod_{t=1}^{T} \mathbb{P}\{D_t = c_t \mid H_t(\overline{c}_{t-1})\} + \prod_{t=1}^{T} \mathbb{P}\{D'_s = c_s \mid H_s(\overline{c}_{s-1})\}$$

$$= \prod_{t \neq s} \mathbb{P}\{D_t = a_t \mid H_t(\overline{a}_{t-1})\} - \prod_{t \neq s} \mathbb{P}\{D'_t = a_t \mid H_t(\overline{a}_{t-1})\}$$

where the first equality follows by the first condition above, the second by adding and subtracting $\big[1 - \mathbb{P}\{D_s = a_s \mid H_s(\overline{a}_{s-1})\}\big]$ and $\big[1 - \mathbb{P}\{D'_s = a_s \mid H_s(\overline{a}_{s-1})\}\big]$, the third by the definition of c_t above, and the final line by the third condition above.

Reaching a contradiction. The analysis above implies that the product of trimming indicators Π_T is equal to the product of propensity scores *without including* the interventional propensity score at timepoint s. However, by assumption we know that trimming occurs at timepoint s, and therefore with positive probability this equality cannot hold, because the right-hand side does not account for trimming at timepoint s. Hence, we have reached a contradiction. □

C Proofs for Section 5

C.1 Helper lemmas of efficient influence function of Q_t

We begin with several general helper lemmas.

Lemma 4. *Under the setup of Proposition 5, $\varphi_m(Z)$ 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 5. *Under the setup of Proposition 5,*

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

and

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

Proof. These follow by iterated expectations on H_t. □

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

Lemma 6. *Under the setup of Proposition 5,*

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

Proof. By adding zero and rearranging,

$$
\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\{\mathbb{P}(A_t = a_t \mid H_t) - \widehat{\mathbb{P}}(A_t = a_t \mid H_t)\Big\} s_t'\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t); k_t\}\{\mathbb{1}(b_t = a_t) - \widehat{\mathbb{P}}(A_t = b_t \mid H_t)\} \\
&\quad + \Big\{\widehat{\mathbb{P}}(A_t = b_t \mid H_t) - \mathbb{P}(A_t = b_t \mid H_t)\Big\}\Big[1 - s_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t); k_t\}\Big] \\
&\quad + \mathbb{1}(b_t = a_t) s_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t); k_t\} + \Big[1 - s_t\{\widehat{\mathbb{P}}(A_t = a_t \mid H_t); k_t\}\Big]\widehat{\mathbb{P}}(A_t = b_t \mid H_t) \\
&\quad - \mathbb{1}(b_t = a_t) s_t\{\mathbb{P}(A_t = a_t \mid H_t); k_t\} - \Big[1 - s_t\{\mathbb{P}(A_t = a_t \mid H_t); k_t\}\Big]\mathbb{P}(A_t = b_t \mid H_t) \\
&\equiv \{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\}\widehat{s}_t'(a_t)\{\mathbb{1}(b_t = a_t) - \widehat{\mathbb{P}}_t(b_t)\} \\
&\quad + \{\mathbb{P}_t(b_t) - \widehat{\mathbb{P}}_t(b_t)\}\{1 - \widehat{s}(a_t)\} \\
&\quad + \mathbb{1}(b_t = a_t)\widehat{s}_t(a_t) + \{1 - \widehat{s}_t(a_t)\}\widehat{\mathbb{P}}_t(b_t) \\
&\quad - \mathbb{1}(b_t = a_t) s_t(a_t) - \{1 - s_t(a_t)\}\mathbb{P}_t(b_t) \\
&= \mathbb{1}(b_t = a_t)\Big[\big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t) + \widehat{s}_t(a_t) - s_t(a_t)\Big] \\
&\quad - \big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t)\widehat{\mathbb{P}}_t(b_t) + \big\{\mathbb{P}_t(b_t) - \widehat{\mathbb{P}}_t(b_t)\big\}\big\{1 - \widehat{s}_t(a_t)\big\} \\
&\quad + \big\{1 - \widehat{s}_t(a_t)\big\}\widehat{\mathbb{P}}_t(b_t) - \{1 - s_t(a_t)\}\mathbb{P}_t(b_t) \\
&= \mathbb{1}(b_t = a_t)\Big[\big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t) + \widehat{s}_t(a_t) - s_t(a_t)\Big] \\
&\quad - \big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}(a_t)\big\}\widehat{s}_t'(a_t)\widehat{\mathbb{P}}_t(b_t) + \mathbb{P}_t(b_t)\{1 - \widehat{s}_t(a_t)\} \\
&\quad - \{1 - s_t(a_t)\}\mathbb{P}_t(b_t) \\
&= \mathbb{1}(b_t = a_t)\Big[\big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t) + \widehat{s}_t(a_t) - s_t(a_t)\Big] \\
&\quad - \{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\}\widehat{s}_t'(a_t)\widehat{\mathbb{P}}_t(b_t) + \mathbb{P}_t(b_t)\{s_t(a_t) - \widehat{s}_t(a_t)\} \\
&= \mathbb{1}(b_t = a_t)\Big[\big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t) + \widehat{s}_t(a_t) - s_t(a_t)\Big] \\
&\quad + \widehat{\mathbb{P}}_t(b_t)\Big[s_t(a_t) - \widehat{s}_t(a_t) - \widehat{s}_t'(a_t)\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\}\Big] \\
&\quad + \{\mathbb{P}_t(b_t) - \widehat{\mathbb{P}}_t(b_t)\}\{s_t(a_t) - \widehat{s}_t(a_t)\}
\end{aligned}
$$

where the first and second lines follows by definition, the third through sixth lines by canceling terms and adding zero.

Second-order Taylor expansions of $s\{\mathbb{P}(A_t = a_t \mid H_t)\}$ yield the result. For the first term in the final display above:

$$
\begin{aligned}
&\mathbb{1}(b_t = a_t)\Big[\big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t) + \widehat{s}_t(a_t) - s_t(a_t)\Big] \\
&= \mathbb{1}(b_t = a_t)\Big(\big\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\big\}\widehat{s}_t'(a_t) + \widehat{s}_t(a_t) - \widehat{s}_t(a_t) - \widehat{s}_t'(a_t)\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\} \\
&\qquad - \frac{1}{2}\widehat{s}_t''(a_t)\big[\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\}^2\big] - o\big[\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\}^2\big]\Big) \\
&= -\mathbb{1}(b_t = a_t)\left(\frac{1}{2}\widehat{s}_t''(a_t)\big[\{\widehat{\mathbb{P}}_t(a_t) - \mathbb{P}_t(a_t)\}^2\big] + o\big[\{\widehat{\mathbb{P}}_t(a_t) - \mathbb{P}_t(a_t)\}^2\big]\right)
\end{aligned}
$$

By essentially the same argument, the second term yields

$$
\widehat{\mathbb{P}}_t(b_t)\Big[s_t(a_t) - \widehat{s}_t(a_t) - \widehat{s}_t'(a_t)\{\mathbb{P}_t(a_t) - \widehat{\mathbb{P}}_t(a_t)\}\Big]
$$

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

from which the result follows. □

C.2 Proposition 5 and Theorem 3

Now, we turn to establishing Proposition 5 and Theorem 3. 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.

C.2.1 Backwards-in-time lemmas

Lemma 7. *Under the setup of Proposition 5,*

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

Proof. We have

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

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

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

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

Lemma 8. *Under the setup of Proposition 5,*

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

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

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 9. *Under the setup of Proposition 5,*

$$\begin{aligned}\mathbb{E}\{\widehat{\varphi}(Z)\} &= m_0 - \widehat{m}_0 \\ &+\sum_{t=0}^{T}\mathbb{E}\left[\left\{\prod_{s=0}^{t}\widehat{r}_s(A_s \mid H_s)-\prod_{s=0}^{t} r_s(A_s \mid H_s)\right\}\{\widetilde{m}_t(A_t, H_t)-\widehat{m}_t(A_t, H_t)\}\right] \\ &+\sum_{t=1}^{T}\mathbb{E}\left[\left\{\prod_{s=1}^{t-1}\widehat{r}_s(A_s \mid H_s)-\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t, H_t)\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}\right] \\ &+\sum_{t=1}^{T}\mathbb{E}\left(\left\{\prod_{s=1}^{t-1} r_s(A_s \mid H_s)\right\}\sum_{b_t}\widehat{m}_t(b_t, H_t)\Big[\mathbb{E}\{\widehat{\phi}_t(b_t; A_t, H_t) \mid H_t\}+\widehat{Q}_t(b_t \mid H_t)-Q_t(b_t \mid H_t)\Big]\right).\end{aligned}$$

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

C.2.2 Forwards-in-time lemmas

Lemma 10. *Under the setup of Proposition 5,*

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

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

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

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

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

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

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

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

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

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

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

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

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

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 11. *Under the setup of Proposition 5,*

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

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

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

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

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 12. *Under the setup of Proposition 5,*

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

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

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

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

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

Proof of Proposition 5

Proof. Lemmas 9, 5, and 6 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. □

C.3 Theorem 3

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 9,

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

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 5 and 6, the triangle inequality, and Cauchy-Schwarz:

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

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

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

Forwards-in-time:
By the same argument above and Lemma 12,

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

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 5 and 6, 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

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

□

C.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 3. 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. □

C.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, $\mathbb{E}\left\{\phi_{t+1}(b_{t+1}; A_{t+1}, H_{t+1}) \mid H_{t+1}\right\} = 0$ by Lemma 5. 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

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

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

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

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

C.6 Theorem 4

Proof. We omit arguments throughout. We have

$$\mathbb{E}\left(\widehat{\psi}^* - \psi\right) = \mathbb{E}\left(\widehat{m}_0^* - m_0\right)$$

$$= \mathbb{E}\left(\widehat{P}_1^*(Z) - m_0\right)$$

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

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 5 and 6, 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

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

□

C.7 Corollary 2

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