Research Article

Hao Sun, Ashkan Ertefaie, Xin Lu, and Brent A. Johnson*

Improved Doubly Robust Estimation in Marginal Mean Models for Dynamic Regimes

Abstract: Doubly robust (DR) estimators are an important class of statistics derived from a theory of semiparametric efficiency. They have become a popular tool in causal inference, including applications to dynamic treatment regimes. The doubly robust estimators for the mean response to a dynamic treatment regime may be conceived through the augmented inverse probability weighted (AIPW) estimating function, defined as the sum of the inverse probability weighted (IPW) estimating function and an augmentation term. The IPW estimating function of the causal estimand via marginal structural model is defined as the complete-case score function for those subjects whose treatment sequence is consistent with the dynamic regime in question divided by the probability of observing the treatment sequence given the subject's treatment and covariate histories. The augmentation term is derived by projecting the IPW estimating function onto the nuisance tangent space and has mean-zero under the truth. The IPW estimator of the causal estimand is consistent if (i) the treatment assignment mechanism is correctly modeled and the AIPW estimator is consistent if either (i) is true or (ii) nested functions of intermediate and final outcomes are correctly modeled.

Hence, the AIPW estimator is doubly robust and, moreover, the AIPW is semiparametric efficient if both (i) and (ii) are true simultaneously. Unfortunately, DR estimators can be inferior when either (i) or (ii) is true and the other false. In this case, the misspecified parts of the model can have a detrimental effect on the variance of the DR estimator. We propose an improved DR estimator of causal estimand in dynamic treatment regimes through a technique originally developed by [4] which aims to mitigate the ill-effects of model misspecification through a constrained optimization.

In addition to solving a doubly robust system of equations, the improved DR estimator simultaneously minimizes the asymptotic variance of the estimator under a correctly specified treatment assignment mechanism but misspecification of intermediate and final outcome models. We illustrate the desirable operating characteristics of the estimator through Monte Carlo studies and apply the methods to data from a randomized study of integrilin therapy for patients undergoing coronary stent implantation. The methods proposed here are new and may be used to further improve personalized medicine, in general.

Keywords: causal inference, informative eligibility, missing data, treatment competing events

2020 Mathematics Subject Classification: 92B15

1 Introduction

Estimating population parameters in the presence of missing data is a common but challenging problem when one desires robust precise estimates under a missing at random assumption [*e.g.* 22]. Horvitz-Thompson

***Corresponding Author: Brent A. Johnson:** Department of Biostatistics and Computational Biology, University of Rochester, Rochester, 14642, New York, United States of America; Email: brent_johnson@urmc.rochester.edu
Hao Sun: Uber Technologies Inc, San Francisco, 94103, California, United States of America
Ashkan Ertefaie: Department of Biostatistics and Computational Biology, University of Rochester, Rochester, 14642, New York, United States of America
Xin Lu: Department of Biostatistics and Programming, Sanofi, Bridgewater, 07657, New Jersey, United States of America

[1952] estimators, defined by multiplying the outcome by the reciprocal of the missingness probability and also called inverse-probability weighted (IPW) estimators, are popular because they are straightforward to implement and consistent if the missingness probability is modeled correctly. At the same time, Horvitz-Thompson estimators can be imprecise and sensitive to estimated missingness probabilities that are very close to zero. Augmented inverse probability weighted (AIPW) estimating functions [16] are constructed as the sum of the IPW estimating function and a mean-zero augmentation term. The optimal augmentation is the orthogonal projection of the IPW estimating function onto the nuisance tangent space [2, 26] and is a function of one or more regression functions, *i.e.* mean outcomes modeled as a function of covariates and also called outcome regression (OR) models. The AIPW estimators are doubly robust (DR) which implies they are consistent whether the missingness probability is modeled correctly or the OR models in the augmentation term are modeled correctly. Moreover, if both models are correctly specified, the AIPW is semiparametric efficient [2, 16, 22]. However, this adaptive estimator is not optimal when only some models are correctly specified [*e.g.* 9, 19, 27].

In this paper, our contribution is to propose a coefficient estimator of the causal parameters in a marginal model for dynamic regimes that has better operating characteristics than the doubly robust estimator proposed by [11]. In particular, we seek to define an estimator that has two properties:

(P1) doubly robust;
(P2) minimum asymptotic variance within a class of augmented estimator assuming the PS model is correct, but the assumed OR model may or may not be correct.

The first optimal doubly robust estimators for missing data problems satisfying properties (P1)-(P2) was Tan's [2006] restricted maximum likelihood estimator (MLE) and the targeted MLE by [27]. Later, [4] proposed a competing estimator that had better finite sample performance than the estimator by [19]. Other improved DR estimators available in the literature include the augmented restricted MLE [21], the augmented OR estimator [17], numerically-derived locally efficient estimators [5], and an adaptation of Godambe's [1960] optimal estimating equation for missing data [13]. Most of the aforementioned techniques demonstrate the principles of their method in the problem of estimating mean outcome when the outcome may be missing due to measured covariates or, analogously, estimating the average causal effect from observational data. These principles can, in theory, be extended to new problems but the details for any given strategy are a substantial challenge when working with complex temporal data.

Our proposed estimator is developed by applying a constrained augmentation technique [4] to the DR estimator in Murphy *et al.* The challenge of this extension is deriving the correct expression for the constrained augmentation. Although both our estimator and the estimator by Murphy *et al.* possess the double robustness property in (P1), our proposed estimator has minimum asymptotic variance when the OR models are incorrect in property (P2) while the estimator by Murphy *et al.* does not. The reason property (P2) is important for this problem, in particular, is because it is extremely difficult to specify OR models that satisfy the nested constraints induced by the marginal model much less to expect that they are specified correctly (See the constraints in (6) and our discussion in Section 3.3). Thus, the proposed method is a hedge against imprecise parameter estimation due to OR model misspecification, which seems more likely than not in the case of marginal models for dynamic regimes.

The remainder of this paper is organized as follows. Section 2 reviews DR estimation of causal estimands in dynamic treatment regimes via marginal models [11] whereas Section 3 details our new methodological contribution. In Section 4, we demonstrate the operating characteristics of our methods through simulation studies and illustrate the utility of our methods through application to data from an infusion study conducted at Duke University. Proofs of technical results are relegated to the Web Appendix.

2 Doubly Robust Estimation in Marginal Models for Dynamic Regimes

2.1 Notation and Assumptions

Without loss of generality, we follow much the same notation given in [11]. Let A_j, $j = 1, \ldots, K$, be the stochastic treatment decision at time t_j and $\overline{A}_j = (A_1, \ldots, A_j)$ be the history of treatment decisions through time t_j. Similarly, define tailoring and auxiliary variables, S_j and X_j, respectively, for $j = 0, \ldots, K - 1$, as well as their histories $\overline{S}_j = (S_0, \ldots, S_j)$ and $\overline{X}_j = (X_0, \ldots, X_j)$. Let $\mathcal{G}$ denote subgroups of interest. The set of all possible potential outcomes is $\{Y(\overline{a}_K) \mid \overline{a}_K \in \overline{\mathcal{A}}_K\}$, where $\overline{\mathcal{A}}_K$ is the collection of all possible treatment allocation vectors; similarly, the set of potential intermediate outcomes is $\{\overline{S}_K(\overline{a}_K) \mid \overline{a}_K \in \overline{\mathcal{A}}_K\}$. Hence, we assume that the potential outcome does not depend on the treatment assignment mechanism nor is affected by others' treatment [14, 18]. We also assume a sequential randomization assumption [14], that is, the treatment assignment A_j is conditionally independent of $\{\mathcal{G}, S_0, \ldots, S_K(\overline{a}_{K-1}), Y(\overline{a}_{K-1})\}$ given $\{L_0, A_1, L_1, \ldots, A_{j-1}, L_{j-1}\}$, with $L_j = (S_j, X_j)$. In words, the sequential randomization implies (i) that at no point in time does treatment assignment depend on future outcomes, and (ii) that the history of auxiliary variables $\overline{X}_{j-1}$ is sufficiently rich to ensure that treatment assignment A_j does not vary systematically by intermediate or potential outcomes within levels of $(\overline{A}_{j-1}, \overline{L}_{j-1})$. The observed data is $O = \{\mathcal{G}, L_0, A_1, L_1, A_2, \ldots, L_{K-1}, A_K, Y\}$, where $Y = Y(\overline{A}_K)$, $S_j = S_j(\overline{A}_j)$, $X_j = X_j(\overline{A}_j)$ for all $j = 1, \ldots, K$.

Let $d_j(\overline{L}_{j-1})$ be a treatment decision or treatment assignment rule at time t_j, $j = 1, \ldots, K$, and the sequence of decision rules for the entire treatment allocation period defines the dynamic treatment regimen, i.e. $\overline{d}_K = (d_1, \ldots, d_K)$. Now, define the product

$$W_{\overline{d}_j}(\overline{A}_j, \overline{L}_{j-1}; \overline{\pi}_{j0}) = \prod_{m=1}^{j} \frac{I\{A_m = d_m(\overline{L}_{m-1})\}}{\pi_{m0}(A_m \mid \overline{A}_{m-1}, \overline{L}_{m-1})}, \quad (j = 1, \ldots, K), \tag{1}$$

where, at time t_m, $\pi_{m0}(\cdot \mid \overline{A}_{m-1}, \overline{L}_{m-1})$ is the treatment assignment probability in the observational data and treatment assignment in the dynamic regime is degenerate according to the treatment decision rule $d_m(\overline{L}_{m-1})$. Without loss of generality, we define $\overline{\pi}_j = (\pi_1, \ldots, \pi_j)$ and the vector of true treatment assignment probabilities $\overline{\pi}_{j0} = (\pi_{10}, \ldots, \pi_{j0})$ for $j = 1, \ldots, K$. The product in (1) plays a critical role in our ability to estimate causal estimands from the observed data and is a version of the non-stabilized inverse PS weight. Let $P_{\overline{d}_K}$ and $E_{\overline{d}_K}$ be probability distribution and expectation under the dynamic regime, respectively, and P and E be the probability distribution and expectation in the observed data, respectively. Then, if there is non-trivial probability that a randomly selected subject can follow the regime $\overline{d}_K$ in the observational, then by Lemma 4.1 of [11], the distribution of $(Y, \overline{S}_{K-1}, \overline{A}_K, \mathcal{G})$ under $P_{\overline{d}_K}$ is absolutely continuous with respect to the distribution of $(Y, \overline{S}_{K-1}, \overline{A}_K, \mathcal{G})$ under P and a version of the Radon-Nikodym derivative is

$$E\left[W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_{K0}) \mid Y = y, \overline{S}_{K-1} = \overline{s}_{K-1}, \overline{A}_K = \overline{a}_K, \mathcal{G} = g\right].$$

2.2 Ordinary doubly robust estimation

Suppose that we are interested in estimating the parameter vector β from the marginal model,

$$E_{\overline{d}_K}(Y \mid \mathcal{G}) = \mu(\beta, \mathcal{G}), \tag{2}$$

where $\mu(\beta, \mathcal{G})$ is a parameterization of the conditional mean of Y given subgroups $\mathcal{G}$ under the dynamic regime $\overline{d}_K$. Then, under regularity assumptions outlined in Section 2.1 as well as in Lemma 4.1 of [11], the statistic

$$W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_{K0})\dot{\mu}(\beta, \mathcal{G})(Y - \mu(\beta, \mathcal{G})),$$

where $\dot{\mu}(\beta, \mathcal{G}) = (\partial/\partial\beta)\mu(\beta, \mathcal{G})$, has mean zero at the true marginal parameter $\beta = \beta_0$. If the treatment assignment probabilities $\{\pi_{j0}\}$ were known *a priori*, this statistic could be used to define an estimator for β.

However, in observational studies, the treatment assignment probabilities are unknown *a priori* and instead one posits a statistical model. To this end, we model the treatment assignment probabilities through a finite-dimensional vector of parameters γ and estimate it through the estimating function,

$$U_\gamma(O; \gamma) = \sum_{j=1}^{K} \sum_{a_j} I(A_j = a_j) u_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma) \tag{3}$$

where $u_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma) = \dot{\pi}_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma) \mathcal{V}_j^{-1}(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)\{a_j - \pi_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)\}$, $\dot{\pi}_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma) = (\partial/\partial\gamma)\pi_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)$ and $\mathcal{V}_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)$ is the conditional variance of a_j given $\overline{A}_{j-1}$ and $\overline{L}_{j-1}$. The estimating function $U_\gamma(O; \gamma)$ may be regarded as proportional to the first-order partial derivative of the log likelihood with respect to γ for the treatment assignment mechanism. As such, $EU_\gamma(O; \gamma_0) = 0$ when the treatment assignment probabilities are correctly modeled. Let $\widehat{\gamma}$ denote the maximum likelihood estimator of γ_0, *i.e.* the solution to the estimating equations, $0 = \mathbb{P}_n U_\gamma(O; \gamma)$. Then, the inverse probability weighted estimator (IPW) for β is $\widehat{\beta}_{\text{IPW}}$, the solution to the system of estimating equations,

$$0 = \mathbb{P}_n \left[\psi_{\text{IPW}}(O; \beta, \overline{\pi}_K(\widehat{\gamma})) \right],$$

where

$$\psi_{\text{IPW}}(O; \beta, \overline{\pi}_K(\gamma)) = W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_K(\gamma))\dot{\mu}(\beta, \mathcal{G})(Y - \mu(\beta, \mathcal{G})), \tag{4}$$

$\overline{\pi}_j(\gamma) = (\pi_1(\gamma), \dots, \pi_j(\gamma))$ and $\pi_j(\gamma) = \pi_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma))$ for $j = 1, \dots, K$. The IPW estimator $\widehat{\beta}_{\text{IPW}}$ is consistent as long as the treatment selection probabilities $\{\pi_j\}$ are correctly modeled such that derivative is consistently estimated, *i.e.* $W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_{K0}) = W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_K(\gamma_0))$.

    A concern of IPW estimators, assuming $\{\pi_j\}$ are correctly modeled, is their efficiency. Briefly, the likelihood may be written as the product of two expressions, $\mathcal{L}(O) = \mathcal{L}_{\text{OR}}(O)\mathcal{L}_{\text{PS}}(O)$, where $\mathcal{L}_{\text{PS}}(O) = \prod_{j=1}^{K} \pi_{j0}(A_j \mid \overline{A}_{j-1}, \overline{L}_{j-1})$ and $\mathcal{L}_{\text{OR}}(O)$ is the product of conditional densities, $\prod_{j=1}^{K} P(L_j^\star \in dL_j \mid \overline{A}_{j-1}, \overline{L}_{j-1})$, where $Y \equiv L_K$ and $P(L_j^\star \in B \mid \cdot) = P(L_j \in B \mid \cdot)$ for every measureable rectangle B. The causal estimand $E_{\overline{d}_K}(Y \mid \mathcal{G})$ is a function of the conditional densities in $\mathcal{L}_{\text{OR}}$ but not $\{\pi_{j0}\}$ in $\mathcal{L}_{\text{PS}}$; hence, the treatment selection probabilities are nuisance parameters for the estimand of interest. The efficient estimator for β will be orthogonal to the score function for the nuisance parameters and, hence, the IPW estimator can be improved by subtracting from ψ_{IPW} its projection onto the score function of the treatment selection probabilities. The projection of ψ_{IPW} onto the score function of the treatment selection probabilities [15] is

$$\mathcal{A}(O; \beta, \overline{\pi}_{K0}, \overline{g}_{K0}\}) = \sum_{j=1}^{K} W_{\overline{d}_j}(\overline{A}_j, \overline{L}_{j-1}; \overline{\pi}_{j0})\dot{\mu}(\beta, \mathcal{G}) \left\{ g_{j0}(\overline{A}_j, \overline{L}_{j-1}) - \mu(\beta, \mathcal{G}) \right\} \tag{5}$$

$$- \sum_{j=1}^{K} \left[\sum_{a_j} \pi_{j0}(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}) \times W_{\overline{d}_j}(a_j, \overline{A}_{j-1}, \overline{L}_{j-1}; \overline{\pi}_{j0})\dot{\mu}(\beta, \mathcal{G})\{g_{j0}(a_j, \overline{A}_{j-1}, \overline{L}_{j-1}) - \mu(\beta, \mathcal{G})\} \right],$$

where the regression functions are nested within one another through the following relationship: $g_{K0}(\overline{A}_K, \overline{L}_{K-1}) = E(Y \mid \overline{A}_K, \overline{L}_{K-1})$ and for $j = 1, \dots, K-1$,

$$g_{j0}(\overline{A}_j, \overline{L}_{j-1}) = E\left[\sum_{a_{j+1}} I\{a_{j+1} = d_{j+1}(\overline{L}_j)\}g_{j+1,0}(a_{j+1}, \overline{A}_j, \overline{L}_j) \mid \overline{A}_j, \overline{L}_{j-1} \right]. \tag{6}$$

Again, we adopt the notation $\overline{g}_j = (g_1, \dots, g_j)$ and the history of true regression models is $\overline{g}_{j0} = (g_{10}, \dots, g_{j0})$. In semi-parametric theory for missing data problems [22, 26], the augmentation term $\mathcal{A}(O; \beta, \overline{\pi}_{K0}, \overline{g}_{K0})$ in (5) has mean zero by construction when evaluated at $\beta = \beta_0$, and the treatment assignment probabilities are modeled correctly so that $\{\pi_{j0}\} \equiv \{\pi_j(\gamma_0)\}$.

    In order to use the augmentation in practice, the true regression functions $\{g_{j0}\}$ must be known or modeled. Often, $g_{j0}(\overline{A}, \overline{L}_{j-1})$ is modeled parametrically or semi-parametrically as $g_j(\overline{A}_j, \overline{L}_{j-1}; \alpha)$ through the finite

dimensional parameter α. For example, Murphy *et al.* [2001] suggested the semi-parametric estimator $\widehat{\alpha}_{\text{MLR}}$, the solution to the estimating equations, $0 = \mathbb{P}_n[U_\alpha(O; \alpha)]$,

$$U_\alpha(O; \alpha) = \dot{g}_K(\overline{A}_K, \overline{L}_{K-1}; \alpha)(Y - g_K(\overline{A}_K, \overline{L}_{K-1}; \alpha)) + \sum_{j=1}^{K} \dot{g}_j(\overline{A}_j, \overline{L}_{j-1}; \alpha) \tag{7}$$

$$\times \left[\sum_{a_{j+1}} I\{a_{j+1} = d_{j+1}(\overline{L}_j)\} g_{j+1}(a_{j+1}, \overline{A}_j, \overline{L}_j; \alpha) - g_j(\overline{A}_j, \overline{L}_{j-1}; \alpha) \right],$$

where $\dot{g}_j(\overline{A}_j, \overline{L}_{j-1}; \alpha) = (\partial/\partial\alpha)g_j(\overline{A}_j, \overline{L}_{j-1}; \alpha)$ for $j = 1, \ldots, K$. Using $\widehat{\alpha}_{\text{MLR}}$ defined through (7), then the usual augmented inverse probability weighted (AIPW) estimator, $\widehat{\beta}_{\text{USUAL}}$, is defined as the solution to the system of equations

$$0 = \mathbb{P}_n \left[\psi_{\text{AIPW}}(O; \beta, \overline{\pi}_K(\widehat{\gamma}), \overline{g}_K(\widehat{\alpha}_{\text{MLR}})) \right],$$

where

$$\psi_{\text{AIPW}}(O; \beta, \overline{\pi}_K(\gamma), \overline{g}_K(\alpha)) = \psi_{\text{IPW}}(O; \beta, \overline{\pi}_K(\gamma)) - \mathcal{A}(O; \beta, \overline{\pi}_K(\gamma), \overline{g}_K(\alpha)),$$

and $\{g_j(\alpha)\} = \{g_j(\overline{A}_j, \overline{L}_{j-1}; \alpha)\}$. The AIPW estimator has the double robustness property: that is, $\widehat{\beta}_{\text{USUAL}}$ is consistent if one of either the treatment selection probabilities $\{\pi_j\}$ or the regression functions $\{g_j\}$ are modeled correctly. If both functions are modeled correctly, the AIPW estimator is semi-parametric efficient [16, 22].

3 New methods

When only one set of functions $\{\pi_j\}$ or $\{g_j\}$ are modeled correctly, the incorrect model can have an adverse effect on the precision of $\widehat{\beta}_{\text{USUAL}}$ unless steps are taken to mitigate the ill-effects of model misspecification. Without loss of generality, suppose that the PS model $\{\pi_j\}$ is correctly modeled but $\{g_j\}$ may or may not be. Under the true $\{g_{j0}\}$, $EU_\alpha(O; \alpha_0) = 0$ and $\widehat{\alpha}_{\text{MLR}} \to_p \alpha_0$. Under incorrectly modeled $\{g_j\}$, however, $\widehat{\alpha}_{\text{MLR}} \to_p \alpha_\star$, for some $\alpha_\star \neq \alpha_0$. Even though the usual estimator $\widehat{\beta}_{\text{USUAL}}$ is consistent regardless of whether $\{g_j\}$ is correctly modeled or not, the variance of $\widehat{\beta}_{\text{USUAL}}$ is only optimal when $\{g_j\}$ is modeled correctly. Thus, we aim to construct an estimator of β that will:

(P1) be doubly robust;
(P2) have smallest asymptotic variance among the class of AIPW estimators when the PS is modeled correctly regardless of whether $\{g_j\}_{j=1}^K$ are modeled incorrectly.

3.1 New Estimation with $\{\pi_j\}$ as Known Functions

In this subsection, we consider constrained doubly robust estimation of β when $\{\pi_j\}$ are known functions of $(\overline{A}_{j-1}, \overline{L}_{j-1})$, *i.e.*, when $\gamma = \gamma_0$ is known. To begin, consider the coefficient estimator $\tilde{\beta}$ defined as the solution to

$$0 = \mathbb{P}_n[\psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))], \tag{8}$$

where $\alpha^\star$ is the limiting value of the estimator $\widehat{\alpha}$ regardless of whether $\{g_j\}$ are modeled correctly or not. The asymptotic variance of $n^{1/2}(\tilde{\beta} - \beta_0)$ is

$$\tilde{V}(\alpha_\star) = \Gamma^{-1} \left\{ \text{var} \lim_{n \to \infty} n^{1/2} \mathbb{P}_n[\psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))] \right\} (\Gamma^{-1})^\top,$$

where Γ is the asymptotic slope matrix of the right-hand side of (8) with respect to β_0, say,

$$\Gamma = E \left\{ \frac{\partial}{\partial\beta} \psi_{\text{AIPW}}(O; \beta, \overline{\pi}_{K0}, \overline{g}_K(\alpha)) \bigg|_{\beta=\beta_0, \alpha=\alpha_\star} \right\}.$$

The asymptotic slope matrix is invariant to the choice of $\alpha_\star$ either directly or in the limit. Therefore, when the $\{\pi_{j0}\}$ are modeled correctly, minimizing the asymptotic variance $\tilde{V}(\alpha_\star)$ reduces to a problem of minimizing $\text{var}[\psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))]$ as a function in $\alpha_\star$. [For a related problem, see expression (4) on p.8 of [23]].

Now, because $\psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))$ has mean zero when evaluated at the true parameters, then minimizing $\text{var}[\psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))]$ is equivalent to minimizing

$$E\left[\psi_{\text{AIPW}}\left(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star)\right)^2\right], \tag{9}$$

where $\alpha_\star$ is the probabilistic limit of some estimator $\hat{\alpha}$. Define α_{opt} as the minimizer of (9) as a function in $\alpha_\star$. Note that α_{opt} also satisfies the following system of equations,

$$0 = E\left[\dot{\mathcal{A}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))\psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_{K0}, \overline{g}_K(\alpha_\star))\right], \tag{10}$$

where

$$\dot{\mathcal{A}}(O; \beta, \overline{\pi}_K, \overline{g}_K(\alpha_\star)) = \sum_{j=1}^{K} \sum_{a_j} \{I(A_j = a_j) - \pi_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1})\} \times W_{\overline{d}_j}(a_j, \overline{A}_{j-1}, \overline{L}_{j-1}; \overline{\pi}_j)\dot{\mu}(\beta, \mathcal{G})\dot{g}_j(\overline{A}_j, \overline{L}_{j-1}; \alpha_\star).$$

When $\{g_j\}$ are correctly specified, $\alpha_{\text{opt}} = \alpha_0$; otherwise, α_{opt} is some other value in the parameter space. To satisfy (P2), we wish to define an estimator $\hat{\alpha}_{\text{opt}} \rightarrow_p \alpha_{\text{opt}}$ while, at the same time, $\hat{\alpha}_{\text{opt}} \rightarrow_p \alpha_{\text{opt}} \equiv \alpha_0$ when $\{g_j\}$ are correctly specified, so that (P1) is satisfied.

To proceed with construction of the improved DR estimator via constrained augmentation, it is convenient to re-express (10). Using Lemmas 1-3 in Appendix A, we show that the estimating function on the right-hand side of (10) can be written as

$$E\left[\{Y - \mu(\beta_0, \mathcal{G})\} \times \sum_{j=1}^{K} \frac{1 - \pi_{j0}(d_j(\overline{L}_{j-1}) \mid \overline{A}_{j-1}, \overline{L}_{j-1})}{\prod_{m=1}^{j} \pi_{m0}(A_m \mid \overline{A}_{m-1}, \overline{L}_{m-1})} \dot{\mu}^2(\beta_0, \mathcal{G})\dot{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha_\star)\right. \tag{11}$$

$$- \sum_{j=1}^{K} \left[\{g_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha_\star) - \mu(\beta_0, \mathcal{G})\} \right.$$

$$\left.\times \left\{ \frac{1 - \pi_{j0}(d_j(\overline{L}_{j-1}) \mid \overline{A}_{j-1}, \overline{L}_{j-1})}{\prod_{m=1}^{j} \pi_{m0}(A_m \mid \overline{A}_{m-1}, \overline{L}_{m-1})} \dot{\mu}^2(\beta_0, \mathcal{G})\dot{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha_\star) \right\} \right]$$

We can then use (11) to propose an estimator for α. We propose to estimate α by solving the system of estimating equations,

$$0 = \mathbb{P}_n \left[\sum_{j=1}^{K-1} W_{\overline{d}_j}\left(\overline{A}_j, \overline{L}_{j-1}; \overline{\pi}_{j0}\right) \dot{\mu}^2(\beta_0, \mathcal{G})q_j(\overline{L}_{j-1}; \overline{\pi}_{j0}, \alpha) \right. \tag{12}$$

$$\times \left\{ g_{j+1}(d_{j+1}(\overline{L}_j), \overline{A}_j, \overline{L}_j; \alpha) - g_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha) \right\} + W_{\overline{d}_K}\left(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_{K0}\right) \dot{\mu}^2(\beta_0, \mathcal{G})q_K(\overline{L}_{K-1}; \overline{\pi}_{K0}, \alpha)$$

$$\times \left\{ Y - g_K(d_K(\overline{L}_{K-1}), \overline{A}_{K-1}, \overline{L}_{K-1}; \alpha) \right\} \right],$$

where, for $r = 1, \ldots, K$,

$$q_r(\overline{L}_{r-1}; \overline{\pi}_r, \alpha) = \sum_{j=1}^{r} \frac{1 - \pi_j(d_j(\overline{L}_{j-1}) \mid \overline{A}_{j-1}, \overline{L}_{j-1})}{\prod_{m=1}^{j} \pi_m(A_m \mid \overline{A}_{m-1}, \overline{L}_{m-1})} \dot{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha). \tag{13}$$

In Appendix A, we show that the estimating function on the right-hand side of (12) has mean zero when one of either $\{\pi_j\}$ or $\{g_j\}$ are correctly specified and, hence, the double robustness property (P1) is satisfied. Furthermore, when the PS model is correctly specified and (12) is evaluated with $\{\pi_{j0}\}$, we show that (12) has mean zero at $\alpha = \alpha_{\text{opt}}$, and therefore, the constrained augmentation improved doubly robust (IDR) estimator

$\widehat{\beta}_{\mathrm{IDR}}$ has the smallest asymptotic variance among the class of AIPW estimators with correctly specified $\{\pi_j\}$, where $\widehat{\beta}_{\mathrm{IDR}}$ is the solution to the AIPW estimating equations with $\widehat{\alpha}$ solving (12). Thus, both properties (P1) and (P2) are satisfied.

Remark 1. The right-hand side of the proposed estimating equation in (12) is a function of the unknown causal estimand β_0. In practice, when the $\{\pi_j\}$ are known, the unknown parameters $(\beta_0, \alpha_\star)$ are estimated jointly by augmenting the estimating equation in (12) with the AIPW estimating equation in (8). The root-solving method one uses in practice may depend, in part, by the complexity of the marginal mean $\mu(\beta_0, \mathcal{G})$ as a function in β_0 and $\mathcal{G}$ in (12). An alternative asymptotically-equivalent practical solution would be to replace β_0 in (12) with a consistent estimator $\widehat{\beta}$ of β_0, *e.g.*, the usual AIPW estimator, and then estimate $(\beta_0, \alpha_\star)$ together by solving the pair of estimating equations (12) and (8).

3.2 New Estimation with $\{\pi_j\}$ as Functions of Unknown γ

We now consider the case where the probabilities $\{\pi_j\}$ are functions of an unknown parameter vector γ. In general, $\mathrm{var}[\psi_{\mathrm{AIPW}}(O; \beta, \overline{\pi}_K(\gamma_0), \overline{g}_K(\alpha)]$ is not equal to $\mathrm{var}[\psi_{\mathrm{AIPW}}(O; \beta, \overline{\pi}_K(\widehat{\gamma}), \overline{g}_K(\alpha)]$ in either finite or large samples. Exceptions include when $\{\pi_j\}$ are known by design or when the estimators $\widehat{\gamma}$ are constructed to be asymptotically orthogonal. The details given in this subsection are appropriate for a majority of cases where there is a non-negligible effect on $\mathrm{var}[\psi_{\mathrm{AIPW}}(O; \beta, \overline{\pi}_K(\gamma_0), \overline{g}_K(\alpha)]$ when γ_0 is replaced with a consistent estimator $\widehat{\gamma}$. When the variance cost to the estimating function ψ_{AIPW} is zero or asymptotically negligible, methods described in Section 3.1 are appropriate exactly or in the limit.

For $j = 1, \ldots, K$, let $\pi_j(\gamma) = \pi_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)$ and $\overline{\pi}_j(\gamma) = \{\pi_1(\gamma), \ldots, \pi_j(\gamma)\}$. Similar to the outline in Section 3.1, we aim to define an estimator $\widehat{\beta}$ that solves the AIPW estimating equations,

$$0 = \mathbb{P}_n \left[\psi_{\mathrm{AIPW}}(O; \beta, \overline{\pi}_K(\widehat{\gamma}), \overline{g}_K(\widehat{\alpha})) \right], \tag{14}$$

where $\widehat{\gamma}$ and $\widehat{\alpha}$ are estimators of γ and α, respectively, such that the double robustness property in (P1) and optimality property in (P2) are both satisfied. From Theorem 9.1 in [22], when the PS model is correctly specified, the AIPW estimating equations in (14) are asymptotically equivalent to

$$0 = \mathbb{P}_n \left[\psi_{\mathrm{AIPW}}(O; \beta, \overline{\pi}_K(\gamma_0), \overline{g}_K(\alpha_\star)) - H_0^\top(\beta, \gamma_0, \alpha_\star) \mathcal{I}_\gamma^{-1} U_\gamma(O; \gamma_0) \right], \tag{15}$$

where $\alpha_\star$ is the limiting value of some estimator $\widehat{\alpha}$ and

$$H_0(\beta, \gamma_0, \alpha_\star) = E \left\{ \left. \frac{\partial}{\partial \gamma} \psi_{\mathrm{AIPW}}(O; \beta, \overline{\pi}_K(\gamma), \overline{g}_K(\alpha)) \right|_{\beta=\beta_0, \gamma=\gamma_0, \alpha=\alpha_\star} \right\},$$

$$\mathcal{I}_\gamma = E \left\{ U_\gamma(O; \gamma_0) U_\gamma^\top(O; \gamma_0) \right\}.$$

Let $c_\star$ be the value of c that minimizes

$$E \left[\psi_{\mathrm{AIPW}} \left(O; \beta_0, \overline{\pi}_K(\gamma_0), \overline{g}_K(\alpha) \right) - c^\top U_\gamma(O; \gamma_0) \right]^2 \tag{16}$$

when $\alpha_\star$ is substituted for α. Our objective is thus to find ζ_{opt} that minimizes (16) where $\zeta = (a, c)$, and the procedure is analogous to finding α_{opt} that solves (10) in Section 3.1. However, in the current scenario, we allow for the fact that $\{\pi_{j0}\}$ are functions of the unknown parameter γ that is estimated from the data. We can then use this correspondence to Section 3.1 to propose an estimator of ζ_{opt}, say $\widehat{\zeta}_{\mathrm{opt}}$, that will in turn lead to the improved doubly robust (IDR) estimator $\widehat{\beta}_{\mathrm{IDR}}$ by solving (14) with smallest asymptotic variance when the PS models are correctly specified but the OR models may or may not be; see Appendix B for a detailed derivation of this section.

Building on the principles and logic laid out above, we therefore propose to estimate ζ by solving the following estimating equation,

$$
0 = \mathbb{P}_n\Bigg[\sum_{j=1}^{K-1} W_{\overline{d}_j}(\overline{A}_j, \overline{L}_{j-1}; \overline{\pi}_j(\widehat{\gamma}))\dot{\mu}^2(\beta, \mathcal{G})\tilde{q}_j(\overline{L}_{j-1}; \zeta, \widehat{\gamma}) \tag{17}
$$

$$
\times \{\tilde{g}_{j+1}(d_{j+1}(\overline{L}_j), \overline{A}_j, \overline{L}_j; \zeta, \widehat{\gamma}) - \tilde{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \widehat{\gamma})\}\Bigg]
$$

$$
+ W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_K(\widehat{\gamma}))\dot{\mu}^2(\beta, \mathcal{G})\tilde{q}_K(\overline{L}_{K-1}; \zeta, \widehat{\gamma})
$$

$$
\times \{Y - \tilde{g}_K(d_K(\overline{L}_{K-1}), \overline{A}_{K-1}, \overline{L}_{K-1}; \zeta, \widehat{\gamma})\}\Bigg],
$$

where

$$
\tilde{q}_r(\overline{L}_{r-1}; \zeta, \gamma) = \sum_{j=1}^{r} \frac{1 - \pi_j(d_j(\overline{L}_{j-1}) \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)}{\prod_{m=1}^{j} \pi_m(A_m \mid \overline{A}_{m-1}, \overline{L}_{m-1}; \gamma)} \begin{bmatrix} \tilde{g}_{j\alpha}(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma) \\ \tilde{g}_{jc}(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma) \end{bmatrix},
$$

and

$$
\tilde{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma) = g_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha) + c^\top \tilde{g}_{jc}(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma),
$$

$$
\tilde{g}_{jc}(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma) =
$$

$$
- W_{\overline{d}_{j-1}}^{-1}(\overline{A}_{j-1}, \overline{L}_{j-2}; \overline{\pi}_{j-1}(\gamma)) \frac{\pi_j(A_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma) \sum_{a_j} I(A_j = a_j) u_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)}{\{I(A_j = d_j(\overline{L}_{j-1})) - \pi_j(d_j(\overline{L}_{j-1}) \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)\}},
$$

and $\tilde{g}_{j\alpha}(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma) = \dot{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \alpha)$, and $u_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}; \gamma)$ is defined in (3). We observe that $\widehat{\gamma}$ converges to γ_0 when the PS models $\{\pi_j\}$ are correctly specified regardless of whether the OR models $\{g_j\}$ are correctly specified or not. Using similar arguments in Appendix A that show the results in Section 3.1, $\widehat{\zeta}_{\text{opt}}$ solving the system of estimating equations in (17), will converges to ζ_{opt} and thus satisfying (P2). On the other hand, when the PS models are misspecified but the OR models are correct, the expectation of (17) can be shown to be zero when $\zeta \equiv (\alpha, c) = (\alpha_0, 0)$ and therefore the double-robustness property (P1) is satisfied. Thus, our estimator $\widehat{\beta}_{\text{IDR}}$, defined as the solution to $0 = \psi_{\text{AIPW}}(O; \beta_0, \overline{\pi}_K(\widehat{\gamma}), \overline{g}_K(\widehat{\alpha}_{\text{opt}}))$ where $\widehat{\alpha}_{\text{opt}}$ is defined through $\widehat{\zeta}_{\text{opt}}$ via (17) and $\widehat{\gamma}$ is an M-estimator defined in (3), is DR and has smallest asymptotic variance among all DR estimators when the PS models are correct but OR models may not be.

We may use a theory of Z-estimation to describe the asymptotic properties of $\widehat{\beta}_{\text{IDR}}$. Specifically, we combine the estimating equations for $\theta = (\beta^\top, \gamma^\top, \alpha^\top)^\top$ and derive the asymptotic properties for $\widehat{\beta}$ from the asymptotic properties of $\widehat{\theta} = (\widehat{\beta}^\top, \widehat{\gamma}^\top, \widehat{\alpha}^\top)^\top$. Combining the estimating functions for β and γ with the proposed estimating function (17) with respect to $\zeta = (\alpha, c)$, $\widehat{\theta}$ solves the system of estimating equations,

$$
0 = \mathbb{P}_n \begin{bmatrix} \psi_{\text{AIPW}}(O; \beta, \overline{\pi}_K(\gamma), \overline{g}_K(\alpha)) \\ U_\gamma(O; \gamma) \\ \begin{bmatrix} \sum_{j=1}^{K-1} W_{\overline{d}_j}(\overline{A}_j, \overline{L}_{j-1}; \overline{\pi}_j(\gamma))\dot{\mu}^2(\beta, \mathcal{G})\tilde{q}_j(\overline{L}_{j-1}; \zeta, \gamma) \\ \times \{\tilde{g}_{j+1}(d_{j+1}(\overline{L}_j), \overline{A}_j, \overline{L}_j; \zeta, \gamma) - \tilde{g}_j(d_j(\overline{L}_{j-1}), \overline{A}_{j-1}, \overline{L}_{j-1}; \zeta, \gamma)\} \\ + W_{\overline{d}_K}(\overline{A}_K, \overline{L}_{K-1}; \overline{\pi}_K(\gamma))\dot{\mu}^2(\beta, \mathcal{G})\tilde{q}_K(\overline{L}_{K-1}; \zeta, \gamma) \\ \times \{Y - \tilde{g}_K(d_K(\overline{L}_{K-1}), \overline{A}_{K-1}, \overline{L}_{K-1}; \zeta, \gamma)\} \end{bmatrix} \end{bmatrix}.
$$

Under conditions used to derive the improved DR estimator [*e.g.*, 11, Appendix], $\widehat{\theta}$ is a Z-estimator with an asymptotic normal distribution whose covariance follows standard formulae [*e.g.* 3, Ch. 7] and can be estimated in various ways, including direct evaluation of the empirical covariance matrix and resampling methods.

3.3 Caveats

The technique described above is general and applicable for all non-random dynamic treatment regimes via the marginal model in (2). This development applies to the work by [4] and [23] who detailed the idea for a population mean in the presence of missing outcome and monotone coarsening, respectively.

Although the principles behind locally efficient doubly robust estimation is technically sound, putting the principles into practice can be a substantial challenge. Firstly, for K-stage trials with K even moderately large and non-trivial treatment effect, there are many regression functions to specify. Such bookkeeping can be an onerous task and encourages the use of a subclass of $\{g_j\}$ that restrict regression coefficients to be common across stages. Second, the regression functions $\{g_j\}$ are highly constrained through their nested definition in (6). This caveat led [11] to propose that $\{g_j\}$ depend on auxiliary variables $\overline{X}_{j-1}$ but not on treatment $\overline{A}_{j-1}$, *i.e.* no treatment effect. Unfortunately, if there is a treatment effect, then one is relying entirely on a correct PS model via $\{\pi_j\}$ for consistency of the AIPW estimator. As we demonstrated above, even if the PS model is correct, incorrectly specified $\{g_j\}$ could have undesirable consequences on the precision of the AIPW estimator and downstream statistical inference. Thus, efficiency of the AIPW estimator in the current context is both germane and important.

4 Examples

4.1 The Setup: Adaptive Treatment Length Studies

Our numerical and real-data examples are illustrations of our continued work in dynamic regimes for infusion studies, where a clinical goal of interest is to estimate the mean outcome if an attending physician stops the infusion at a time of their choice or at the time of an unexpected random event (*i.e.* a time not of their choice), whichever time comes first [8]. In these applications, as with many dynamic treatment regimes, evaluating eligibility criteria for treatment assignment is a key feature of the analysis. Eligibility for treatment assignment is evaluated sequentially at each stage or clinic visit at time t_j and depends on treatment history, covariate history, and the history of eligibility. In the infusion study, treatment assignment at t_j is synonymous with a physician's decision to stop or continue treatment by choice at time t_j provided the patient satisfies eligibility criteria. In the observational study, treatment assignment depends on treatment history, covariate history, and the history of eligibility whereas treatment assignment depends on treatment and eligibility history in the non-random dynamic regime for our infusion study application.

To define the decision rule leading to our dynamic regime, let a target treatment length t_r be given, $r \in \{1, \ldots, K\}$. Let S_j be a binary tailoring variable. We consider the decision rule

$$d_j(\overline{L}_{j-1}) = \begin{cases} \text{stop treatment,} & \text{if } t_j = t_r \text{ and } \overline{S}_{j-1} = 0; \\ \text{continue,} & \text{otherwise.} \end{cases} \tag{18}$$

In the infusion study application, a subject initiates infusion treatment at t_0. According to the decision rule in (18), at t_j, $j < r$, the attending physician continues treatment as long as the subject satisfies the eligibility criteria, *i.e.* $\overline{S}_j = 0$, where S_j is a random (in)eligibility event in the interval $(t_j, t_{j+1}]$. If Y is the clinical endpoint of interest in the infusion study, then our goal is to estimate the causal estimand, $E_{\overline{d}_K}(Y \mid \mathcal{G}) = \beta$, using the data.

The observed data for the proposed application is outlined in Table 1. In contrast to a dynamic regime whose treatment length is stopped at t_r provided $\overline{S}_{r-1} = 0$, subjects may stop at any t_j, $j < K$, in the observational study. In addition, the concept of eligibility for treatment assignment at t_j is broader in the observational study. In order to be eligible for treatment assignment at t_j, a subject must have not stopped treatment at any prior time point $t_{j'}$, $j' < j$, as well as avoided random adverse events $\{S_{j'} = 1\}$ for all $j' < j$. In our notation, $A_j = 1$ implies a decision to stop or "assign" treatment at t_j by choice and the probability that a

Table 1: The observed data in adaptive treatment length studies. Time t_j^+ denotes the moment just after t_j.

	Innovation at t_j	Description
Treatment assignment, A_j	$A_j I(\overline{A}_{j-1} = 0, \overline{S}_{j-1} = 0)$	stop or continue at t_j given subject is eligible at t_j;
History of covariates and tailoring variables, $\overline{L}_j(\overline{A}_j)$	$(\overline{X}_j, S_j) I(\overline{A}_j = 0, \overline{S}_{j-1} = 0)$	covariate history and ineligibility indicator in $(t_j, t_{j+1}]$ given subject eligible at t_j^+;
	$I(\overline{A}_j = 0, \overline{S}_{j-1} = 0)$	indicator that subject is eligible for treatment assignment at t_j^+;
Endpoint, $Y(\overline{A}_K)$	NA	outcome measured at end of study.

provider stops treatment at t_j given a subject is eligible at t_j is

$$\lambda_j(\overline{X}_{j-1}) = P(A_j = 1 \mid \overline{X}_{j-1}, \overline{A}_{j-1} = 0, \overline{S}_{j-1} = 0).$$

Then, the treatment assignment mechanism is modeled as

$$\pi_j(a_j \mid \overline{A}_{j-1}, \overline{L}_{j-1}) = \left\{ \lambda_j(\bar{X}_{j-1}) \right\}^{a_j} \left\{ 1 - \lambda_j(\bar{X}_{j-1}) \right\}^{I(\overline{A}_j = 0, \overline{S}_{j-1} = 0)}. \tag{19}$$

Adhering to the non-stochastic dynamic regime that defines the adaptive treatment length policy is given by the degenerate treatment assignment mechanism,

$$I\{a_j = d_j(\overline{L}_{j-1})\} = I(t_j = t_r)^{a_j} \times \{1 - I(\overline{A}_r = 0, \overline{S}_{r-1} = 0)\}^{I(\overline{A}_j = 0, \overline{S}_{j-1} = 0)}.$$

Combining the degenerate treatment assignment mechanism above with the mechanism in (19) from the observed data leads to the definition of the Radon-Nikodym derivative, *i.e.*,

$$W_{\overline{d}_j}(\overline{A}_j, \overline{L}_{j-1}) = \prod_{m=1}^{j} \frac{I\{A_m = d_m(\overline{L}_{m-1})\}}{\pi_m(A_m \mid \overline{A}_{m-1}, \overline{L}_{m-1})}$$

$$= \prod_{m=1}^{j} \left\{ \frac{I(t_m = t_r)}{\lambda_m(\bar{X}_{m-1})} \right\}^{A_m} \left\{ \frac{1 - I(\overline{A}_r = 0, \overline{S}_{r-1} = 0)}{1 - \lambda_m(\bar{X}_{m-1})} \right\}^{I(\overline{A}_m = 0, \overline{S}_{m-1} = 0)},$$

that is critical for estimating the causal estimand $E_{\overline{d}_K}[Y \mid \mathcal{G}]$ from observational data.

4.2 Simulation Studies

We performed simulation studies to evaluate the finite sample performance of competing estimators of for the causal parameter β in the marginal model $\mu(\beta, \mathcal{G})$, including outcome regression (OR), inverse probability weighted (IPW), ordinary double robust and improved double robust estimators. First, we partition covariate history as $\overline{X}_j = ((V_0^{(1)})^\top, (\overline{V}_j^{(2)})^\top)^\top$, where $V_0^{(1)}$ baseline (time-invariant) covariates, and $\overline{V}_j^{(2)} = (V_0^{(2)}, V_1^{(2)}, \ldots, V_j^{(2)})$ is the history of time-varying covariate information, respectively. Next, treatment assignment probabilities are generated as Bernoulli random variables with success probability $\lambda_j(\overline{X}_{j-1}; \gamma) = H(\gamma_j + \gamma_X^\top \overline{X}_{j-1})$, $j = 1, \ldots, K$, where $H(z) = 1/\{1 + \exp(-z)\}$ and $\gamma_X = (\gamma_{V^{(1)}}^\top, \gamma_{V^{(2)}}^\top)$. Intermediate outcomes S_j are similarly simulated as Bernoulli random variables with success probability $\lambda_j^\xi(\overline{X}_j; \alpha_0) = P(S_j = 1 \mid \overline{X}_j, \overline{A}_{j-1} = 0, \overline{S}_{j-1} = 0)$, $j = 0, \ldots, K-1$, where $\lambda_j^\xi(\overline{X}_j; \alpha) = H(\alpha_{\xi,j} + \alpha_{\xi X}^\top \overline{X}_j)$ and $\alpha_{\xi X} = (\alpha_{\xi,V^{(1)}}^\top, \alpha_{\xi,V^{(2)}}^\top)^\top$. Finally, the endpoint Y is assumed to follow a linear model, $Y = g_{K0}(\overline{A}_K, \overline{L}_{K-1}; \alpha) + \epsilon$, where $g_{K0}(\overline{A}_K, \overline{L}_{K-1}; \alpha_0) = E(Y \mid \overline{A}_K, \overline{L}_{K-1})$,

$$g_{K0}(\overline{A}_K, \overline{L}_{K-1}; \alpha) = \alpha_{y0} + \sum_{j=1}^{K} \alpha_{yj} I(\{A_j = 1\} \cup \{S_{j-1} = 1\}) + \sum_{j=1}^{K} \alpha_{yS_{j-1}} I\{S_{j-1} = 1\} + \alpha_{yX}^\top \overline{X}_{K-1},$$

$\alpha_{yX} = (\alpha_{yV^{(1)}}^{\top}, \alpha_{yV^{(2)}}^{\top})^{\top}$, and ϵ is an independent standard normal random variable. The parameters in the conditional mean model $g_{K0}(\overline{A}_K, \overline{L}_{K-1})$ have practical interpretation for our application. The parameters α_{yj} describe the adjusted effect of treatment length on the outcome Y irregardless of the reason why treatment was stopped. Recall, treatment starts at t_0 and must end in one of the intervals, $(t_{j-1}, t_j]$, $j = 1, \ldots, K$. The parameters α_{yS_j} allow for the possibility that the presence or absence of the intermediate events $(S_0, \ldots, S_{K-1})$ modifies the adjusted effect of treatment length on mean outcome. The parameters α_{yX} describe the adjusted effect of covariate history on mean outcome. An overview of the simulation scheme used to generate data in Table 1 is outlined in Algorithm 1.

We generated data according to the aforementioned scenario for $K = 2$ stages, with two baseline covariates $V_0^{(1)}$ and history of time-varying covariate information defined as its current value, *i.e.* $\overline{V}_j^{(2)} = V_j^{(2)}$. The parameter vector for treatment assignment is $\gamma = (\gamma_1, \gamma_{V^{(1)}}^{\top}, \gamma_{V^{(2)}})^{\top}$, $\gamma_1 = 10/7$, $\gamma_{V^{(1)}} = (-1/5, -1/5)^{\top}$, and $\gamma_{V^{(2)}} = -4/5$. The parameter vector for intermediate outcomes is $\alpha_{\xi} = (\alpha_{\xi,0}, \alpha_{\xi,1}, \alpha_{\xi,V^{(1)}}^{\top}, \alpha_{\xi,V^{(2)}})^{\top}$ with $\alpha_{\xi,0} = 13/7$, $\alpha_{\xi,1} = 26/7$, $\alpha_{\xi,V^{(1)}} = (-1/10, -1/10)^{\top}$, and $\alpha_{\xi,V^{(2)}} = -9/10$. The outcome parameter vector is $\alpha_y = (\alpha_{y0}, \alpha_{y1}, \alpha_{y2}, \alpha_{yV^{(1)}}^{\top}, \alpha_{yV^{(2)}})^{\top}$, with $(\alpha_{y0}, \alpha_{y1}, \alpha_{y2}) = (3/2, 1/2, 3/10)$, $\alpha_{V^{(1)}} = (1, 1/2)^{\top}$, and $\alpha_{V^{(2)}} = 1/2$ or 1. For the particular simulation results presented below, we set $\alpha_{yS_j} = 0$ which assumes that the adjusted effect of treatment length on mean outcome is not modified by the reason for stopping treatment. True values for causal estimands are computed by numerical integration via Monte Carlo method.

Algorithm 1 Pseudocode for simulated data

1: Simulate baseline covariate information: $V_0^{(1)} \sim N_2(0, I_2)$
2: Set $j = 0$
3: **while** $j < K$ **do**
4:      Simulate time-dependent covariate information: if $j = 0$, $V_0^{(2)} \sim N(4, 1)$, else $V_j^{(2)} \sim N(V_{j-1}^{(2)}, 1)$
5:      Simulate tailoring variable: $S_j \sim \text{Bern}\{\lambda_j^{\xi}(\overline{X}_j; \alpha)\}$, where $\lambda_j^{\xi}(\overline{X}_j; \alpha) = H(\alpha_{\xi,j} + \alpha_{\xi,V^{(1)}}^{\top} V_0^{(1)} + \alpha_{\xi,V^{(2)}} V_j^{(2)})$
6:      **if** $s_j = 1$ **then**
7:          Infusion stopped in $(t_j, t_{j+1}]$ due to terminating event; Simulate outcome $Y = \alpha_{y0} + \alpha_{y(j+1)} + \alpha_{yS_j} + \alpha_{yV^{(1)}}^{\top} V_0^{(1)} + \alpha_{yV^{(2)}} V_j^{(2)} + \epsilon$; Stop
8:      **if** $j = K - 1$ **then**
9:          Set $a_{j+1} = K$; Simulate outcome $Y = \alpha_{y0} + \alpha_{yK} + \alpha_{yV^{(1)}}^{\top} V_0^{(1)} + \alpha_{yV^{(2)}} V_j^{(2)} + \epsilon$; Stop
10:      **else**
11:          Simulate treatment assignment at t_{j+1}: $A_{j+1} \sim \text{Bern}\{\lambda_{j+1}(\overline{X}_j; \gamma)\}$, where $\lambda_j(\overline{X}_j; \gamma) = H(\gamma_j + \gamma_{V^{(1)}}^{\top} V_0^{(1)} + \gamma_{V^{(2)}} V_j^{(2)})$
12:          **if** $a_{j+1} = 1$ **then**
13:              Infusion stopped at t_j; Simulate outcome $Y = \alpha_{y0} + \alpha_{y(j+1)} + \alpha_{yV^{(1)}}^{\top} V_0^{(1)} + \alpha_{yV^{(2)}} V_j^{(2)} + \epsilon$; Stop
14:          **else** Increment $j = j + 1$; Continue

In addition to studying the simulation scenario where all models are correctly specified, we also studied scenarios when one or both of the propensity score (PS) or outcome regression (OR) models are incorrectly specified. Specifically, we adopt the following:

- (PS incorrect): In the treatment assignment model, the true success probability of the Bernoulli random variable is modeled as a quadratic function in the time-dependent covariate $V^{(2)}$, *i.e.*, $\lambda_{j0}(\overline{X}_{j-1}; \gamma) = H\{\gamma_j + \gamma_{V^{(1)}}^{\top} V_0^{(1)} + \gamma_{V^{(2)}} (V_j^{(2)})^2\}$, where $\gamma_{V^{(2)}} = 0.2$, but the procedure incorrectly fits the model $\lambda_j(\overline{X}_{j-1}; \gamma) = H(\gamma_j + \gamma_{V^{(1)}}^{\top} V_0^{(1)} + \gamma_{V^{(2)}} V_j^{(2)})$;
- (OR incorrect): We misspecify the model on Y by ignoring the time-dependent covariate $V_j^{(2)}$ in and modeling baseline covariates $V_0^{(1)}$ only.

Note that the true causal estimands remain the same under potentially misspecified PS or OR models when the regression parameter for the time-dependent covariate $V_j^{(2)}$, i.e. $\alpha_{V^{(2)}}$, is fixed at $1/2$ or 1.

A total of five competing estimators are compared under each scenario. The inverse probability weighted (IPW) and three variations of doubly robust estimators rely on a correctly specified treatment assignment mechanism for optimal performance. Here, the treatment assignment probabilities are estimated via maximum likelihood, i.e., $\widehat{\gamma} = -\text{argmin}_\gamma \log \mathcal{L}_{\mathbf{PS}}(O; \gamma)$,

$$\log \mathcal{L}_{\mathbf{PS}}(O; \gamma) = \mathbb{P}_n[\sum_{j=1}^K A_j \log \lambda_j(\overline{X}_{j-1}; \gamma) + I(\overline{A}_j = 0, \overline{S}_{j-1} = 0)\log\{1 - \lambda_j(\overline{X}_{j-1}; \gamma)\}],$$

In addition to the ordinary double robust (USUAL) and improved doubly robust (IDR) estimators, we also computed a naïve improved doubly robust estimator, that ignores the variation in the estimator $\widehat{\gamma}$ in optimizing α, i.e. $\widehat{\beta}_{\mathbf{IDR}_{naive}}$ estimates α by minimizing the L_2-norm $\mathbb{P}_n\{\psi_{\mathbf{AIPW}}(O; \beta, \gamma_0, \alpha)\}^2$. An outcome regression (OR) estimator is computed as

$$\widehat{\beta}_{\mathbf{OR}} = \mathbb{P}_n\left[\sum_{a_1} I\{a_1 = d_1(L_0)\}g_1(a_1, L_0; \widehat{\alpha})\right],$$

where $\widehat{\alpha} = (\widehat{\alpha}_\gamma^\top, \widehat{\alpha}_\xi^\top)^\top$. These estimators are compared in Monte Carlo studies through the following statistics: Monte Carlo bias and standard deviation (SSE), the Monte Carlo average of the sample standard errors (SEE), and empirical coverage probability (ECP) based on a Wald-type 95% confidence interval. Simulation results for a sample size of $n = 300$ are presented in Table 2.

When both PS and OR models are correctly specified, all estimators have relatively small bias and approximately correct coverage probabilities. The OR estimators have smaller finite sample variance than IPW and doubly robust (DR) estimators while the IPW estimator is least precise among all estimators, about 20% larger than any of the DR estimators. When the PS model is incorrect but the OR model is correct, the IPW estimator exhibits substantial finite sample bias while the other estimators do not. When the PS model is correct but the OR model is incorrect, the OR estimator exhibits substantial finite sample bias while the other estimators do not. These finite sample results agree with theoretical results and is the primary motivation for doubly robust estimators.

The usual and improved doubly robust (IDR) estimators have similar performance when the OR model is correct. The main advantage of the IDR estimator compared to the usual DR estimator occurs when the PS model is correct but the OR model is incorrect. In this case, we see that Monte Carlo standard deviation of the usual DR estimator is 6% larger when $\alpha_{V^{(2)}} = 0.5$ and 37% larger when $\alpha_{V^{(2)}} = 1$. Moreover, the simulation results when $\alpha_{V^{(2)}} = 1$ show clearly that the naïve IDR estimator yields suboptimal estimates of the causal estimand even though the performance seems very similar to IDR when $\alpha_{V^{(2)}} = 0.5$. The rationale for suboptimal performance, in general, is that choosing $\widehat{\alpha}$ to minimize the variance of the AIPW estimator when the PS is correct and γ_0 known is not the same as the variance of the AIPW estimator when PS is correct but γ_0 is unknown and must be estimated [4, 19, 21, 23, 27]. Thus, because α is not estimated optimally in the naïve IDR procedure, i.e. $\widehat{\alpha}$ does not converge in probability to α_{opt}, the naïve IDR estimator of the causal estimand is not locally efficient.

Additionally, in our simulation studies, the DR estimators had considerably better performance than the OR estimator even when both the PS and OR models were incorrect. Similar results have been shown by other authors [4, 23] and illustrate sensitivity of the DR estimators and scale of the bias when these estimators are not guaranteed to be consistent. In short, OR estimators are superior when the OR model is specified correctly. But if one is interested in guarding against possible misspecification of the OR model, doubly robust estimators are worth the effort.

As discussed in Section 3, the asymptotic covariance of $\widehat{\beta}_{\mathbf{IDR}}$ can be estimated by the empirical sandwich matrix. Alternatively, one can use resampling or perturbation methods to estimate the sampling variation of the estimator and can be justified along the lines of Lu and Johnson [10, Appendix]. As we demonstrate in Table 2, the empirical variance of the bootstrap resamples approximates well the sampling variation of the estimator and results in confidence intervals with approximately nominal coverage.

Table 2: Simulation results for estimating the mean potential outcome in a dynamic treatment regime. Table entries include Monte Carlo bias, simulation standard error (SSE), standard error estimate (SEE), and empirical coverage probability (ECP) for a 95% confidence interval.

$\alpha_{V^{(2)}}$		Bias	SSE	SEE	ECP	Bias	SSE	SEE	ECP
0.5		PS correct, OR correct				PS incorrect, OR correct			
	IPW	−0.013	0.247	0.238	0.920	−0.103	0.306	0.253	0.842
	OR	−0.003	0.114	0.110	0.934	−0.014	0.114	0.111	0.946
	USUAL	0.003	0.192	0.193	0.962	−0.005	0.206	0.209	0.946
	IDR_{naive}	0.005	0.194	0.194	0.956	0.013	0.204	0.214	0.956
	IDR	0.006	0.193	0.196	0.954	0.017	0.207	0.220	0.960
		PS correct, OR incorrect				PS incorrect, OR incorrect			
	IPW	−0.012	0.247	0.238	0.920	−0.102	0.306	0.253	0.842
	OR	−0.263	0.111	0.110	0.360	−0.295	0.114	0.110	0.248
	USUAL	−0.001	0.207	0.216	0.936	−0.096	0.264	0.241	0.850
	IDR_{naive}	0.002	0.195	0.198	0.950	−0.003	0.209	0.229	0.960
	IDR	0.005	0.196	0.200	0.956	0.035	0.217	0.237	0.950
1.0		PS correct, OR correct				PS incorrect, OR correct			
	IPW	−0.014	0.299	0.276	0.920	−0.183	0.340	0.282	0.770
	OR	−0.008	0.125	0.122	0.954	−0.009	0.126	0.124	0.942
	USUAL	0.018	0.203	0.202	0.950	0.021	0.211	0.204	0.944
	IDR_{naive}	0.017	0.202	0.202	0.944	0.038	0.219	0.217	0.936
	IDR	0.019	0.202	0.206	0.948	0.040	0.220	0.226	0.954
		PS correct, OR incorrect				PS incorrect, OR incorrect			
	IPW	−0.018	0.299	0.276	0.920	−0.187	0.340	0.282	0.770
	OR	−0.530	0.133	0.132	0.014	−0.573	0.135	0.132	0.008
	USUAL	−0.001	0.288	0.293	0.926	−0.172	0.379	0.297	0.750
	IDR_{naive}	0.006	0.226	0.224	0.962	0.003	0.236	0.248	0.956
	IDR	0.013	0.210	0.218	0.954	0.071	0.241	0.259	0.932

4.3 Data Analysis

We also demonstrate the utility of the methods through an analysis of data from the Enhanced Suppression of the Platelet IIb/IIIa Receptor with Integrilin Therapy (ESPRIT) trial [12]. Briefly, ESPRIT was designed to investigate a novel dosing regimen of eptifibatide infusion in patients undergoing elective stent percutaneous coronary intervention (PCI). For those patients randomized to the experimental infusion arm, eptifibatide was delivered as two boluses plus an infusion, with infusion starting immediately after the first bolus and continuing until it was stopped by physician discretion or treatment-competing event. Although the protocol provided general recommendations on when providers ought to stop infusion, the decision to stop infusion was ultimately left to physician choice. The protocol also articulated when infusion must be discontinued due to adverse events. Thus, the adaptive treatment length policy treats to the target time t_j or infusion-terminating event, whichever comes first. In the ESPRIT study, however, infusion continues until infusion-terminating event or when the attending physician chooses to discontinue it, whichever come first. The study endpoint is a composite endpoint of death, myocardial infarction (MI), or urgent target revascularization in 30 days. The techniques described above in Section 2, coupled with the details in Section 4.1, allow us to estimate the mean composite endpoint in the adaptive treatment length policy using data from the observed ESPRIT trial data.

Results from our data analysis on 1040 patients receiving the experimental infusion regimen are presented in Table 3. All estimators of causal estimands adjust for potential confounders including angina and weight at baseline as well as time-dependent enzyme level measured during post-surgery observation. Note that the composite endpoint is a Bernoulli outcome but that the IPW estimator will work the same regardless

Table 3: Results from the infusion trial. Table entries are multiplied by 10^4.

	IPW	USUAL	IDR$_{naive}$	IDR
t_1	624 (92)	667 (163)	630 (93)	626 (93)
t_2	866 (128)	864 (128)	870 (128)	877 (128)

of whether the outcome is continuous or discrete; other estimator rely on outcome regression (OR) models via $\{g_j\}$. To exemplify potential advantages of the improved doubly robust estimator, we adopt linear regression models in $\{g_j\}$ thus choosing a potentially poor link function in a generalized linear model for a binary outcome. We computed proportions of the composite endpoint under adaptive treatment length policies at $t_1 = 16$ and $t_2 = 22$ hours. We first observe that the range of estimates across methods is modest: $6.2-6.7\%$ at 16 hours and $8.6-8.7\%$ at 22 hours. However, the range in standard error estimates is huge by comparison. The standard error of the ususal doubly robust estimator is 77% larger than all other estimators, including the inefficient IPW estimator. We conjecture this large standard error of the usual doubly robust estimator is likely due to our choice of link function in $\{g_j\}$. But the improved doubly robust estimators counterbalance, in part, the deficiencies in the outcome regression models and yields a precise estimate of the estimand as the theory suggests ought to happen.

5 Conclusions

In this paper, we proposed an improved doubly robust estimator of causal estimands in marginal models for dynamic treatment regimes [11] through a technique first proposed by [4]. The primary objective of the technique is to improve the precision of the doubly robust estimator when the propensity score (PS) via $\{\pi_j\}$ is correctly specified but the outcome regression (OR) models via $\{g_j\}$, $g_j = g_j(\overline{A}_j, \overline{L}_{j-1}; \alpha)$, may be incorrectly specified. This is achieved by defining a new estimator for the regression parameter α to minimize the asymptotic variance of the modified AIPW estimator via (17).

Our numerical and real data examples in Section 4 illustrate potentially unnerving aspects of AIPW and ordinary doubly robust estimation [19, 20, 27]. While the augmentation in AIPW estimation is intended to improve the precision of the IPW estimator, specifying the OR models may not be straightforward for complex data and incorrectly specifying OR models can have unfortunate consequences. The $\{g_j\}$ for dynamic regimes are examples of complicated OR regression functions and, as shown in Section 4.3, can have a dramatic adverse impact on statistical inference if modeled incorrectly. In our data example in Section 4.3, poor choice of models for $\{g_j\}$ increased standard error estimates for one of the causal estimands by more than 75% compared to other inverse-weighted estimators. Improved doubly robust estimators do not suffer from these deficiencies but require more effort on the part of the data analyst.

As the proposed method aims to improve the efficiency of DR estimators in marginal models for dynamic treatment regimes, there are other methods [*e.g.* 17, 19, 20, 27] that aim to improve the efficiency and robustness of ordinary doubly robust (DR) estimators. It would be interesting to compare some of these competing methods in examples of dynamic treatment regimes, such as the treatment length problem. Such investigation would take considerable programming effort and resources and it is beyond the scope of the current manuscript.

In this paper, we have imposed finite-dimensional parametric models on the nuisance parameters. Alternatively, one could relax this modeling assumption and use data-adaptive techniques to estimate the nuisance parameters. However, in this case, the asymptotic linearity of the resulting estimator will rely on certain rate assumptions for all the nuisance parameter estimators. Targeted learning and undersmoothing of regularization parameters offer a possible remedy to this issue [*e.g.* 1, 24, 25]. Generalizations of these methods to treatment length policy settings merit further research.

References

[1] Benkeser, D., M. Carone, M. J. van der Laan, and P. Gilbert (2017): "Doubly robust nonparametric inference on the average treatment effect," *Biometrika*, 104, 863–880.

[2] Bickel, P., C. Klassen, Y. Ritov, and J. Wellner (1993): "Efficient and adaptive estimation for semiparametric models," *Baltimore: Johns Hopkins University Press*.

[3] Boos, D. D. and L. A. Stefanski (2013): *Essential Statistical Inference*, Springer: New York.

[4] Cao, W., A. A. Tsiatis, and M. Davidian (2009): "Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data," *Biometrika*, 96, 723–734, https://doi: 10.1093/biomet/asp033.

[5] Frangakis, C. E., T. Qian, Z. Wu, and I. Diaz (2015): "Deductive derivation and Turing-computerization of semiparametric efficient estimation," *Biometrics*, 71, 867–874, https://doi:10.1111/biom.12362.

[6] Godambe, V. (1960): "An optimum property of regular maximum likelihood estimation," *Annals of Mathematical Statistics*, 31, 1208–1211.

[7] Horvitz, D. G. and D. J. Thompson (1952): "A generalization of sampling without replacement from a finite universe," *Journal of the American Statistical Association*, 47, 663–685, https://doi:10.1080/01621459.1952.10483446.

[8] Johnson, B. A. and A. A. Tsiatis (2004): "Estimating mean response as a function of treatment duration in an observational study, where duration may be informatively censored," *Biometrics*, 60, 315–323, https://doi:10.1111/j.0006-341X.2004.00175.x.

[9] Kang, D. Y. J. and J. L. Schafer (2007): "Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data (with discussion and rejoinder)," *Statistical Science*, 22, 523–539, https://doi:10.1214/07-STS227.

[10] Lu, X. and B. A. Johnson (2015): "Direct estimation of the mean outcome on treatment when treatment assignment and discontinuation compete," *Biometrika*, 102, 797–807, https://doi:10.1093/biomet/asv043.

[11] Murphy, S. A., M. J. van der Laan, and J. M. Robins (2001): "Marginal mean models for dynamic regimes," *Journal of the American Statistical Association*, 96, 1410–1423, https://doi:10.1198/016214501753382327.

[12] O'Shea, J. C., M. Madan, W. J. Cantor, C. M. Pacchiana, S. Greenberg, D. M. Joseph, M. M. Kitt, T. J. Lorenz, and J. E. Tcheng (2000): "Design and methodology of the esprit trial: evaluating a novel dosing regimen of eptifibatide in percutaneous coronary intervention," *American Heart Journal*, 140, 834–839, https://doi:10.1067/mhj.2000.110458.

[13] Qin, J., B. Zhang, and D. H. Y. Leung (2017): "Efficient augmented inverse probability weighed estimation in missing data problems," *Journal of Business & Economic Statistics*, 35, 86–97, https://doi:10.1080/07350015.2015.1058266.

[14] Robins, J. M. (1997): "Causal inference from complex longitudinal data," in M. Berkane, ed., *Latent Variable Modeling and Applications to Causality*, number 120 in Lecture Notes in Statistics, New York: Springer Verlag, 69–117, https://doi:10.1007/978-1-4612-1842-5_4.

[15] Robins, J. M. (1999): "Testing and estimation of direct effects by reparameterizing directed acyclic graphs with structural nested models," *Computation, Causation, and Discovery*, 349–405.

[16] Robins, J. M., A. Rotnizky, and L. P. Zhao (1994): "Estimation of regression coefficients when some regressors are not always observed," *Journal of the American Statistical Association*, 89, 846–866, https://doi:10.1080/01621459.1994.10476818.

[17] Rotnizky, A., Q. Lei, M. Sued, and J. M. Robins (2012): "Improved double-robust estimation in missing data and causal inference models," *Biometrika*, 99, 439–456, https://doi:10.1093/biomet/ass013.

[18] Rubin, D. B. (1986): "Comment: Which ifs have causal answers," *Journal of the American Statistical Association*, 81, 961–962, https://doi:10.1080/01621459.1986.10478355.

[19] Tan, Z. (2006): "A distributional approach for causal inference using propensity scores," *Journal of the American Statistical Association*, 101, 1619–1637, https://doi:10.1198/016214506000000023.

[20] Tan, Z. (2007): "Understanding OR, PS and DR," *Statistical Science*, 22, 560–568, https://doi:10.1214/07-STS227A.

[21] Tan, Z. (2010): "Bounded, efficient, and doubly robust estimation with inverse weighting," *Biometrika*, 97, 661–682, https://doi:10.1093/biomet/asq035.

[22] Tsiatis, A. A. (2006): *Semiparametric Theory and Missing Data*, Springer.

[23] Tsiatis, A. A., M. Davidian, and W. Cao (2011): "Improved doubly robust estimation when data are monotonely coarsened, with application to longitudinal studies with dropout," *Biometrics*, 67, 536–545, https://doi:10.1111/j.1541-0420.2010.01476.x.

[24] van der Laan, M. J. (2014): "Targeted estimation of nuisance parameters to obtain valid statistical inference," *International Journal of Biostatistics*, 10, 29–57.

[25] van der Laan, M. J., D. Benkeser, and W. Cai (2019): "Efficient estimation of pathwise differentiable target parameters with the undersmoothed highly adaptive lasso," *arXiv preprint arXiv:1908.05607*.

[26] van der Laan, M. J. and J. M. Robins (2003): *Unified methods for censored longitudinal data and causality*, Springer Science & Business Media.

[27] van der Laan, M. J. and D. Rubin (2006): "Targeted maximum likelihood learning," *International Journal of Biostatistics*, 2, https://doi:10.2202/1557-4679.1043.