1. Introduction

Identifying an average causal effect of a treatment with observational data requires adjustment for background variables that affect both the treatment and the outcome under study. Often parametric models are assumed for parts of the joint distribution of the treatment, outcome and background variables (covariates) and large sample properties of estimators are derived under the assumption that the parametric models are correctly specified.

A class of semiparametric estimators called inverse probability weighting (IPW) estimators use the difference of the weighted means of the outcomes for the treatment groups as an estimator of the average causal effect, see e.g. Lunceford and Davidian (2004). IPW estimators reweights the observed outcomes to a random sample of all potential outcomes, missing and observed, by letting each observed outcome account for itself and other individuals with similar characteristics proportionally to the probability of their outcome being observed. IPW estimators are used in applied literature (Kwon, Jeong, et al. 2015) and their properties have been studied in the missing data and causal inference literature, see e.g. Vansteelandt, Carpenter, et al. (2010) and Seaman and White (2013) for reviews.

Earlier research have considered properties of IPW estimators for estimating the average causal effect under the assumption that a parametric propensity score (PS) model is correctly specified (Lunceford and Davidian 2004; Yao, Sun, et al. 2010). Properties of IPW estimators using different weights, often referred to as stabilized (Hernán, Brumback, et al. 2000; Hernán and Robins 2006) or normalized (Hirano and Imbens 2001; Busso, DiNardo, et al. 2014) have been discussed together with the impact of violations to an assumption of overlap (*positivity*) (Khan and Tamer 2010; Petersen, Porter, Gruber, Wang, and van der Laan 2010).

To decrease the reliance on the choice of a parametric model of the PS an approach with doubly or multiply robust estimators have emerged. An estimator is referred to as a doubly robust (DR) estimator (Bang and Robins 2005; Tsiatis 2007) since it is a consistent estimator of the average causal effect if either the model for the propensity score or the outcome regression (OR) model is correct (Scharfstein, Rotnitzky, et al. 1999). The efficiency of the DR estimator is a key property and its variance has been described under correct specification of at least one of the models (Cao, Tsiatis, et al. 2009). When both models are correct the estimator reaches the semiparametric efficiency bound described in Robins and Rotnitzky (1994). The large sample properties of IPW estimators with standard, normalized and variance minimized weights, together with a prototypical DR estimator were studied and compared in Lunceford and Davidian (2004) under correct specification of the PS and OR models.

There are few studies on doubly or multiply robust estimators under misspecification of both the PS and the OR models. Kang and Schafer (2007) studied and compared the performance of various DR and non-DR estimators under misspecification of both models. They concluded that many DR methods perform better than simple inverse probability weighting. However a regression-based estimator under a misspecified model was not improved upon. The paper was commented and the relevance of the results were discussed by several authors see e.g., Tsiatis and Davidian (2007), Tan (2007), Robins, Sued, et al. (2007). In Waernbaum (2012) a matching estimator was compared to IPW and DR estimators under misspecification of both the PS and OR models. Here, a robustness class for the matching estimator under misspecification of the PS model was described.

In this paper we describe three commonly used semi-parametric estimators of the average causal effect under the assumption that none of the working models are correctly specified. For this purpose we study the difference between the probability limit of the estimator under model misspecification and the true average causal effect. The purpose of this definition of the bias is that the estimators under study converges to a well-defined limit however not necessarily consistent for the true average causal effect. We study the biases of two IPW estimators and a DR estimator and compare them under the same misspecification of the PS-model. In the comparisons with the DR estimator the biases provide a means to describe when two wrong models are better than one. To analyze the consequences of the model misspecifications we compare the absolute values of the biases in two parts separately, one for each of the two potential outcome means (μ_1, μ_0). For the comparisons we provide sufficient and necessary conditions for inequalities involving the absolute value of the biases of the different estimators. We use a running example of a data generating process with misspecified PS and OR models to illustrate the inequalities. A simulation study is performed to investigate the biases for finite samples. The data generating processes and the misspecified models from the simulation designs are also used for numerical approximations of the large sample properties derived in the paper.

In recent studies strategies for bias reduction under model misspecification have been proposed by inclusion of additional conditions in the estimating equations for both IPW (Imai and Ratkovic 2014) and DR estimators (Vermeulen and Vansteelandt 2015). However, the general approach for analyzing model misspecification provided in this paper could also be used to study the biased reduced estimators.

The paper proceeds as follows. Section 2 presents the model and theory together with the estimators and their properties when the working models are correctly specified. Section 3 presents a general approach and assumptions to study model misspecification. In Section 4 the generic biases are derived and comparisons between the estimators are performed. We present a simulation study in Section 5 containing both finite sample properties of the estimators and numerical large sample approximations and thereafter we conclude with a discussion.

2. Model and theory

The potential outcome framework defines a causal effect as a comparison of potential outcomes that would be observed under different treatments (Rubin 1974). Let X be a vector of pre-treatment variables, referred to as covariates, T a binary treatment, with realized value $T = 1$ if treated and $T = 0$ if control. The causal effect of the treatment is defined as a contrast between two potential outcomes, for example the difference, $Y(1) - Y(0)$, where $Y(1)$ is the potential outcome under treatment and $Y(0)$ is the potential outcome under the control treatment. The observed outcome Y is assumed to be the potential outcomes for each level of the observed treatment $Y = TY(1) + (1 - T)Y(0)$, so that the data vector that we observe is (T_i, X_i, Y_i), where $i = 1, \ldots, n$ are assumed independent and identically distributed copies. In the remainder of the paper we will drop the subscript i for the random variables when not needed. Since each individual only can be subject to one treatment either $Y(1)$ or $Y(0)$ will be missing. If the treatment is randomized the difference of sample averages of the treated and controls will be an unbiased estimator of the average causal effect $\Delta = E[Y(1) - Y(0)]$, the parameter of interest. In the following we will use the notation $\mu_1 = E[Y(1)]$ and $\mu_0 = E[Y(0)]$. When the treatment is not assigned at random the causal effect of the treatment can be estimated if all confounders are observed

Assumption 1. *[No unmeasured confounding]*
$Y(t) \perp\!\!\!\perp T|X, t = 0, 1.$

and if the treated and controls have overlapping covariate distributions

Assumption 2. *[Overlap]*
$\eta < P(T = 1|X) < 1 - \eta$, *for some* $\eta > 0$.

Throughout the paper we assume that Assumptions 1 and 2 hold. Under these assumptions we can estimate the average causal effect with the observed data by marginalizing over the conditional means

$$\Delta = E\left[E(Y \mid X, T = 1) - E(Y \mid X, T = 0)\right]. \tag{1}$$

For matching/stratification estimators (see Imbens and Wooldridge 2009 for a review) the inner expectation in (1) is evaluated by grouping treated and controls in matched pairs or strata formed by the resulting cells of the cross classification of the covariates. Instead of comparing treated and controls on a high dimensional vector of the covariates it is sufficient to condition on a scalar function of the covariates, $e(X) = P(T = 1|X)$, called the propensity score (Rosenbaum and Rubin 1983). Instead of conditioning on the propensity score as in (1), the propensity score can be used as a weight

$$\Delta = E\left[\frac{TY}{e(X)} - \frac{(1-T)Y}{1-e(X)}\right] = E\left[Y(1) - Y(0)\right],$$

where the last equality follows from Assumption 1.

Usually it is assumed that the propensity score and the outcome regression follow parametric models.

Assumption 3. *[Propensity score model]*
The propensity score $e(X)$ *follows a model* $e(X,\beta)$ *parametrized by,* $\beta = (\beta_1,\ldots,\beta_p)$ *and* $\hat{e}(X)$ *is the estimated propensity score* $e(X,\hat{\beta})$ *with a* $n^{1/2}$*-consistent estimator of* β

Assumption 4. *[Outcome regression model]*
The conditional expectation, $\mu_t(X) = E(Y(t)|X)$*,* $t = 0,1$ *follows a model* $\mu_t(X,\alpha_t)$*,* $t = 0,1$ *parametrized by* $\alpha_t = (\alpha_{t1},\ldots,\alpha_{tq_t})$ *and* $\hat{\mu}_t(X)$ *is the estimated outcome regression* $\mu_t(X,\hat{\alpha}_t)$ *with a* $n^{1/2}$*-consistent estimator of* α_t*.*

Consider the example, $e(X,\beta) = [1+\exp(-X'\beta)]^{-1}$ and $\hat{e}(X)$ are the fitted values of the propensity score when $\hat{\beta}$ is a maximum likelihood estimator of β. Similarly the outcome regression model could be a linear model $\mu_t(X,\alpha_t) = X'\alpha_t$ where $\hat{\mu}_t(X)$, $t = 0,1$ are the fitted values when $\hat{\alpha}_t$ is the ordinary least squares estimator.

We study two IPW estimators and a DR estimator described in Lunceford and Davidian (2004). We denote by $\hat{\Delta}_{\text{IPW}_1}$ an estimator defined by:

$$(2) \qquad \hat{\Delta}_{\text{IPW}_1} = \frac{1}{n}\sum_{i=1}^{n}\frac{T_iY_i}{\hat{e}(X_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-T_i)Y_i}{1-\hat{e}(X_i)}.$$

The variance of $\hat{\Delta}_{\text{IPW}_1}$ is

$$(3) \qquad \sigma^2_{\text{IPW}_1} = V_{\text{IPW}_1} - a^TI^{-1}a$$

where V_{IPW_1} is the asymptotic variance when the propensity score is known

$$V_{\text{IPW}_1} = E\left[\frac{Y(1)^2}{e(X)} + \frac{Y(0)^2}{1-e(X)}\right] - (\mu_1-\mu_0)^2,$$

a is a $(p\times 1)$ vector

$$a = E\left[\ \left(\frac{Y(1)}{e(X)} + \frac{Y(0)}{1-e(X)}\right)e'(X)\ \right]$$

for the partial derivatives $e'(X) = \partial/\partial\beta\,\{e(X,\beta)\}$ and I is the $p\times p$ covariance matrix of the estimated propensity score. In $\hat{\Delta}_{\text{IPW}_1}$ each observed treated individual is weighted by $1/e(X)$ and each control is weighted by $1/[1-e(X)]$. Since the weights generate the missing potential outcomes for each of the treatment and control groups respectively we want to divide the weighted sum with the number of individuals in the generated sample consisting of the observed and missing potential outcomes which may not be n for a given sample (Hirano, Imbens, et al. 2003). This gives an IPW estimator $\hat{\Delta}_{\text{IPW}_2}$ with normalized weights

$$(4) \qquad \hat{\Delta}_{\text{IPW}_2} = \left(\sum_{i=1}^{n}\frac{T_i}{\hat{e}(X_i)}\right)^{-1}\sum_{i=1}^{n}\frac{T_iY_i}{\hat{e}(X_i)} - \left(\sum_{i=1}^{n}\frac{1-T_i}{1-\hat{e}(X_i)}\right)^{-1}\sum_{i=1}^{n}\frac{(1-T_i)Y_i}{1-\hat{e}(X_i)},$$

and variance

$$\sigma^2_{\mathrm{IPW}_2} = V_{\mathrm{IPW}_2} - b^T I^{-1} b \tag{5}$$

where V_{IPW_2} is the asymptotic variance when the propensity score is known

$$V_{\mathrm{IPW}_2} = E\left[\frac{(Y(1)-\mu_1)^2}{e(X)} + \frac{(Y(0)-\mu_0)^2}{1-e(X)}\right],$$

and b is a $(p \times 1)$ vector

$$b = E\left[\ \left(\frac{Y(1)-\mu_1}{e(X)} + \frac{Y(0)-\mu_0}{1-e(X)}\right) e'(X)\ \right].$$

Under Assumptions 1-3 the IPW estimators are consistent estimators of the average causal effect Δ with asymptotic distribution $\sqrt{n}(\hat{\Delta}_{\mathrm{IPW}_k} - \Delta) \sim N(0, \sigma^2{}_{\mathrm{IPW}_k})$, $k = 1, 2$.

In addition we study a DR estimator (Lunceford and Davidian 2004; Tsiatis 2007)

$$\begin{aligned}\hat{\Delta}_{\mathrm{DR}} =& \frac{1}{n}\sum_{i=1}^{n} \frac{T_i Y_i - (T_i - \hat{e}(X_i))\hat{\mu}_1(X_i)}{\hat{e}(X_i)} \\ &- \frac{1}{n}\sum_{i=1}^{n} \frac{(1-T_i) Y_i + (T_i - \hat{e}(X_i))\hat{\mu}_0(X_i)}{1-\hat{e}(X_i)}.\end{aligned} \tag{6}$$

Under Assumptions 1-4 we have the large sample distribution $\sqrt{n}(\hat{\Delta}_{\mathrm{DR}} - \Delta) \sim N(0, \sigma^2_{\mathrm{DR}})$ where

$$\sigma^2_{\mathrm{DR}} = V_{\mathrm{IPW}_2} - d, \tag{7}$$

and

$$d = E\left[\left(\sqrt{\frac{1-e(X)}{e(X)}}\,(\mu_1(X)-\mu_1) + \sqrt{\frac{e(X)}{1-e(X)}}\,(\mu_0(X)-\mu_0)\right)\right]^2,$$

with the property that $\sigma^2_{\mathrm{DR}} \leq \sigma^2_{\mathrm{IPW}_1}, \sigma^2_{\mathrm{IPW}_2}$ which was shown by the theory of Robins and colleagues (Robins and Rotnitzky 1994).

3. Model misspecification: a general approach

Our interest lies in the behaviors of the estimators when the propensity score and the outcome regression models are misspecified. For this purpose we replace Assumptions 3 and 4 with two other assumptions defining the probability limit of the estimators under a general misspecification. The misspecifications will further be used to define a general bias of the IPW and DR-estimators. When the propensity score is misspecified an estimator, e.g., a quasi maximum likelihood estimator (QMLE) is not consistent for β in Assumption 3. However, a probability limit for an estimator under model misspecification exists under general conditions, see e.g. White (1982, Theorem 2.2) for QMLE or Wooldridge (2010, Section 12.1) and Boos and Stefanski (2013, Theorem 7.1) for estimators that can be written as a solution of an estimating equation (M-estimators).

In the following, and as an alternative to Assumptions 3 and 4, we will assume that such limits exists. Below we define an estimator $\hat{e}^*(X)$ of the propensity score under a misspecified model $e^{\text{mis}}(X, \beta^*)$.

Assumption 5. *[Misspecified PS model parameters]*
Let $\hat{\beta}^*$ be an estimator under model misspecification, $e^{mis}(X, \beta^*)$, then $\hat{\beta}^* \xrightarrow{p} \beta^*$.

Under model misspecification the probability limit of $\hat{\beta}^*$ is generally well defined however $e^{\text{mis}}(X, \beta^*)$ is not equal to the propensity score $e(X)$. In the following we use the notation $\hat{e}^*(X) = e^{\text{mis}}(X, \hat{\beta}^*)$ as the estimated propensity score and $e^*(X) = e^{\text{mis}}(X, \beta^*)$ under Assumption 5. Below we give an example for true and misspecified parametric models, however, for Assumption 5 we do not need the existence of a true parametric model.

Example 1 For one confounder X and a true PS model $e(X, \beta) = [1 + \exp(-\beta_0 - \beta_1 X - \beta_2 X^2)]^{-1}$ assume that we misspecify the propensity score with a probit model $e^{\text{mis}}(X, \beta^*) = \Psi(-\beta_0^* - \beta_1^* X)$, i.e., we misspecify the link function and omit a second order term. Let $\hat{\beta}^* = (\hat{\beta}_0^*, \hat{\beta}_1^*)$ be the QMLE estimator of the parameters in $e^{\text{mis}}(X, \beta^*)$ obtained by maximizing the quasi-likelihood

$$\ln \mathcal{L} = \sum_{i=1}^{n} \left(T_i \ln e^{\text{mis}}(X_i, \beta^*) + (1 - T_i) \ln(1 - e^{\text{mis}}(X_i, \beta^*)) \right),$$

Then $\hat{e}^*(X) = \Psi(-\hat{\beta}_0^* - \hat{\beta}_1^* X)$, $\hat{\beta}^* = (\hat{\beta}_0^*, \hat{\beta}_1^*) \xrightarrow{p} \beta^* = (\beta_0^*, \beta_1^*)$ under Assumption 5 and $e^*(X) = \Psi(-\beta_0^* - \beta_1^* X)$.

When considering the existence of true and misspecified parametric models, as illustrated in Example 1, the parameters in β and the limiting parameters β^* under the misspecified model need not to be of the same dimension. For instance, the true model could contain higher order terms and interactions that are not present in the estimation model.

The next assumption concerns overlap under model misspecification.

Assumption 6. *[Overlap under misspecification]*
$\nu < e^*(X) < 1 - \nu$, *for some* $\nu > 0$.

In addition to the PS model we also consider misspecified outcome regression models, $\mu_t^{\text{mis}}(X, \alpha_t^*)$, $t = 0, 1$. Denote by $\hat{\alpha}_t^*$, $t = 0, 1$ the estimator of the parameters in $\mu_t^{\text{mis}}(X, \alpha_t^*)$.

Assumption 7. *[Misspecified OR model parameters]*
Let $\hat{\alpha}_t^*$ be an estimator under model misspecification $\mu_t^{mis}(X, \alpha_t^*)$, $t = 0, 1$, then $\hat{\alpha}_t^* \xrightarrow{p} \alpha_t^*$, $t = 0, 1$.

In the following we use the notation $\hat{\mu}_t^*(X) = \mu_t^{\text{mis}}(X, \hat{\alpha}_t^*)$ as the estimated OR and $\mu_t^*(X) = \mu_t^{\text{mis}}(X, \alpha_t^*)$ under Assumption 7 and μ_t^* for the expected value $E\left[\mu_t^*(X)\right]$, $t = 0, 1$.

Assumptions 5 and 7 are defined for misspecified PS and OR models for the purpose of describing their influence on the estimation of Δ. The estimators (2), (4) and (6) can be written by estimating equations where the equations solving for the PS and OR parameters are set up below the main equation for the IPW and DR estimators, see e.g Lunceford and Davidian (2004) and Williamson, Forbes, and White (2014). Assuming parametric PS and OR models the IPW estimators correspond to solving $2+p$ estimating equations $\sum_{i=1}^{n}\psi(\theta, Y_i, T_i, X_i) = 0$ for the parameters $\theta_{\mathrm{IPW}_k} = (\mu_1, \mu_0, \beta)$, $k = 1, 2$ and for the DR estimator $2+p+q_1+q_0$ estimating equations for the parameters $\theta_{\mathrm{DR}} = (\mu_1, \mu_0, \beta, \alpha_1, \alpha_0)$. Using the notation for the misspecified models in Assumptions 5 and 7 the estimating equations change according to the dimensions of the parameters β^* and α_t^*, $t = 0, 1$. A key condition for Assumptions 5 and and 7 to hold is that the misspecification of the PS and/or OR provides estimating equations that uniquely define the parameter although, as a consequence of the misspecification, it will not be the true average causal effect. In the next section we present the asymptotic bias for the IPW and DR estimators under study with general expressions including the limits of the misspecified propensity score and outcome regression models.

4. Bias resulting from model misspecification

4.1. **General biases.** In order to study the large sample bias of $\hat{\Delta}_{\mathrm{IPW}_1}$, $\hat{\Delta}_{\mathrm{IPW}_2}$ and $\hat{\Delta}_{\mathrm{DR}}$ under model misspecification we define the estimators $\hat{\Delta}^*_{\mathrm{IPW}_1}$, $\hat{\Delta}^*_{\mathrm{IPW}_2}$ and $\hat{\Delta}^*_{\mathrm{DR}}$ by replacing $\hat{e}(X)$ in Equations (2), (4) and (6) with $\hat{e}^*(X)$. For the DR-estimator we additionally replace $\hat{\mu}_t(X)$ with $\hat{\mu}_t^*(X)$, $t = 0, 1$.

To assess the properties of the estimators we assume 1, 2, 5, 6 and 7 and regularity conditions for applying a weak law of large numbers for averages with estimated parameters, see Appendix A.1. Note that Assumption 3 and 4 are no longer needed. We evaluate the difference between the probability limits of the estimators under model misspecification and the average causal effect Δ for the IPW and DR estimators:

Theorem 8 (Bias under model misspecification for $\hat{\Delta}^*_{\mathrm{IPW}_1}$)**.** *Under Assumptions 1-2 and 5-6*

$$\hat{\Delta}^*_{IPW_1} - \Delta \xrightarrow{p} E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right] - (\mu_1 - \mu_0).$$

Theorem 9 (Bias under model misspecification for $\hat{\Delta}^*_{\mathrm{IPW}_2}$)**.** *Under Assumptions 1-2 and 5-6*

$$\hat{\Delta}^*_{IPW_2} - \Delta \xrightarrow{p} \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \frac{E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right]}{E\left[\frac{1-e(X)}{1-e^*(X)}\right]} - (\mu_1 - \mu_0).$$

Theorem 10 (Bias under model misspecification for $\hat{\Delta}^*_{\mathrm{DR}}$). *Under Assumptions 1-2 and 5-7*

$$\hat{\Delta}^*_{DR} - \Delta \xrightarrow{p} E\left[\frac{(e(X)-e^*(X))\,(\mu_1(X)-\mu_1^*(X))}{e^*(X)}\right] + E\left[\frac{[e(X)-e^*(X)]\,(\mu_0(X)-\mu_0^*(X))}{(1-e^*(X))}\right].$$

See Appendix A.2 for proofs.

We refer to the limits in Theorem 8, 9, and 10 as the asymptotic biases of the respective estimators, i.e., $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_1})$, $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_2})$ and $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{DR}})$ although they are the difference between the probablity limits of the estimators and the true Δ and not the difference in expectations. The double robustness property of $\hat{\Delta}^*_{\mathrm{DR}}$ is displayed by Theorem 10 since if either $e(X) = e^*(X)$ or $\mu_t(X) = \mu_t^*(X)$, $t = 0, 1$ we have that $\hat{\Delta}^*_{\mathrm{DR}} \xrightarrow{p} \Delta$.

To provide an illustrative example of the biases of the estimators we obtain the misspecified models' limits by misspecifying the link functions in generalized linear models. However, other data generating processes under Assumptions 1-2, 5-7 could also be used. For the propensity score we use binary response models with logit link (true) and a cloglog link (misspecified), for the outcome regression models we use poisson models with log links (true) and gaussian models (misspecified) with identity links. We use numerical approximations to provide values on the parameters in $e^*(X)$ and $\mu_t^*(X)$, $t = 0, 1$ under the given true and misspecified models $e(X)$, $e^{\mathrm{mis}}(X)$, $\mu_t(X)$ and $\mu_t^{\mathrm{mis}}(X)$, $t = 0, 1$.

Example 2 [Bias from link misspecifications]
Let $X \sim Uniform(-2, 2)$ and $T \sim Bernoulli(e(X))$. Assume that

$$e(X) = [1 + \exp(0.5 - X)]^{-1}, \quad e^*(X) = 1 - \exp[-\exp(-0.81 + 0.74X)],$$
$$\mu_1(X) = \exp(2.3 + 0.14X), \quad \mu_0(X) = \exp(1.4 + 0.20X),$$
$$\mu_1^*(X) = 10.06 + 1.48X, \quad \mu_0^*(X) = 4.14 + 0.79X.$$

The marginal means are $\mu_1 = 10.11$ and $\mu_0 = 4.16$, so that $\Delta = 5.94$. Here we have that, $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_1}) = -0.16$, $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_2}) = 0.05$, and $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{DR}}) = -0.02$

4.2. **Comparisons.** To analyze the impact of the model misspecification on the estimators' biases in Section 4.1 we compare the biases for two parts separately. The first part concerns the bias with respect to μ_1 and the second part with respect to μ_0. The first part of $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_1})$ in Theorem 8 is

$$(8) \qquad E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - \mu_1 = cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1,$$

and the first part of $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_2})$ in Theorem 9 is

$$(9) \qquad \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1 = \frac{cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}.$$

For $\hat{\Delta}^*_{\text{IPW}_1}$ we see in (8) that the mean difference between the expected value of the conditional outcome, scaled with the PS-model error, and the marginal outcome contributes to the bias. For $\hat{\Delta}^*_{\text{IPW}_2}$ the contribution (9) is the difference between the conditional and the marginal outcome with the same error scaling, but here, the expected value of the PS-model error, $E\left[e(X)/e^*(X)\right]$, also enters the bias in the denominator. For $\hat{\Delta}^*_{\text{IPW}_1}$ we see from the right hand side of 8 that the sign depends on the covariance of $[e(X)/e^*(X)]$ and $\mu_1(X)$, and the sign of $E\left[e(X)/e^*(X) - 1\right]\mu_1$. For $\hat{\Delta}^*_{\text{IPW}_2}$ we see from the right hand side of (9) that the sign of depends on the covariance only. Hence, the part of the biases described above can be in different directions for the same model misspecification, see also Example 2 for the bias in total.

In the sequel we will give results concerning the absolute values of the first part of the biases in Theorems 8, 9 and 10 but the results can be directly applied for the second part of the biases by replacing $e(X)/e^*(X)$, with $(1-e(X))/(1-e^*(X))$ and $\mu_1(X)$ with $\mu_0(X)$, see Appendix A.3. We define $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$, $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ as

$$(10) \qquad \text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1}) = E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - \mu_1,$$

$$(11) \qquad \text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2}) = \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1,$$

$$(12) \qquad \text{Bias}_1(\hat{\Delta}^*_{\text{DR}}) = E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\left(\mu_1(X) - \mu_1^*(X)\right)\right].$$

Below, we give a sufficient and necessary condition for when the absolute value of $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$ is smaller than the absolute value of $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$.

Theorem 11 (Comparing $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$)**.**

$$\left|\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1\right| < \left|E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - \mu_1\right|.$$

if and only if

$$\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]\right| < E\left[\frac{e(X)}{e^*(X)}\right]\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1\right|.$$

The proof follows directly from the multiplication rules of absolute values. The theorem displays that the comparison of the biases depends on the magnitude and sign of $E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1$ and the covariance. It is no surprise that the covariance of $e(X)/e^*(X)$

and $\mu_1(X)$ (and similarly of $[1-e(X)]\,/\,[1-e^*(X)]$ and $\mu_0(X)$) plays a role for the bias of the estimators. If $\mu_1(X)$ was a constant it could be taken out of the expectations of the first terms in (10) and (11) and the PS-model error would be cancelled by the denominator $E\left[e(X)/e^*(X)\right]$. In this case the bias in (11) would be 0, and thus smaller than (10).

Example 2 1 [Example 2 revisited for $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$]

$$E\left[\frac{e(X)}{e^*(X)}\right] = 0.98, \quad cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] = 0.065, \quad \mu_1 = 10.11.$$

To apply Theorem 11 for the data generating process in Example 2 we evaluate

$$\begin{aligned}\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]\right| &< E\left[\frac{e(X)}{e^*(X)}\right]\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1\right| \\ |0.064| &< 0.98 \cdot |0.064 - 0.02 \cdot 10.11| \\ &= 0.13,\end{aligned}$$

and we conclude that $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2}) < \text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$.

We proceed by investigating the difference between the bias of the IPW estimators (10), (11) and the bias of the DR estimator (12). Hence, we analyze the question of when two wrong models are better than one. We start by comparing $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$. In the following theorem we show a necessary condition for $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ to be smaller than $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$. In the sequel all proofs are provided in Appendix A.3.

Theorem 12 (Necessary condition for $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ smaller than $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$)**.** *If*

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right| < \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|,$$

then

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right]\right| < 2 \cdot \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|.$$

The theorem states that if the DR estimator improves upon the simple IPW-estimator under misspecification of both the PS and the OR model, then, the absolute value of the misspecified outcome model is less than double the absolute value of the true conditional mean under the same scaling of the PS-model error.

Below we give two examples of sufficient conditions for the DR-estimator to have a smaller bias than the simple IPW estimator.

Theorem 13 (Sufficient conditions for Bias($\hat{\Delta}^*_{\text{DR}}$) smaller than Bias($\hat{\Delta}^*_{\text{IPW}_1}$)). *If a)* $\mu_1^* = \mu_1$ *and*

$$\left|cov\left[\tfrac{e(X)}{e^*(X)}, \mu_1(X)\right] - cov\left[\tfrac{e(X)}{e^*(X)}, \mu_1^*(X)\right]\right| < \left|cov\left[\tfrac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\tfrac{e(X)}{e^*(X)} - 1\right]\mu_1\right|,$$

or,

b)

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right]\right| < 2 \cdot \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|,$$

and $E\left[\left(\tfrac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right]$ *and* $E\left[\left(\tfrac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]$ *are either both positive or both negative, then*

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right| < \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|.$$

One of the criteria in a), that $\mu_1 = \mu_1^*$, is reasonable to assume when the corresponding moment condition is used in the estimation of the misspecified outcome model. Also, we have that criterion b) is the same as the necessary condition with the added assumption that the expectation of the (PS-error scaled) conditional outcomes have the same sign.

Example 2 2 [Example 2 revisited for $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$]
For the data generating process in Example 2 we investigate the first part of the bias

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right| < \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|$$
$$0.01 < 0.11$$

implies

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right]\right| < 2 \cdot \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|$$
$$0.10 < 0.22$$

which is consistent with Theorem 12. Assessing the sufficient conditions, in Theorem 13 a) we have that

$$\mu_1 = 10.11 \approx \mu_1^* = 10.07, \quad cov\left[\frac{e(X)}{e^*(X)}, \mu_1^*(X)\right] = 0.074,$$

and thereby,

$$\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] - cov\left[\frac{e(X)}{e^*(X)}, \mu_1^*(X)\right]\right| < \left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1\right|$$
$$|-0.064 + 0.074| < |-0.064 - 0.02 \cdot 10.11|$$
$$0.01 < 0.27.$$

The conditions in b) cannot be applied since $E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]$ and $cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]$ do not have the same sign.

Theorem 14 (Necessary condition for $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ smaller than $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$)**.** *If*

$$\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)(\mu_1(X)-\mu_1^*(X))\right]\right| < \left|\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}-\mu_1\right|,$$

then

$$E\left[\left(\frac{e(X)}{e^*(X)}-1\right)(\mu_1(X))\right]-\frac{\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]} < E\left[\left(\frac{e(X)}{e^*(X)}-1\right)[\mu_1^*(X)]\right]$$
$$< E\left(\left[\frac{e(X)}{e^*(X)}-1\right][\mu_1(X)]\right)+\frac{\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]}.$$

From the theorem we see that for the DR estimator to improve upon the normalized IPW estimator we need that the outcome misspecification is within an interval defined by the true conditional outcome and the absolute value of the covariance. This means that the smaller the covariance is, the more accuracy of the outcome model is required for the $\hat{\Delta}^*_{\text{DR}}$ to be less biased than $\hat{\Delta}^*_{\text{IPW}_2}$.

Under similar assumptions as in Theorem 13 we describe sufficient conditions for the comparison of $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$.

Theorem 15 (Sufficient conditions for $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ smaller than $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$)**.** *If a)* $\mu_1=\mu_1^*$ *and* $\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]-cov\left[\frac{e(X)}{e^*(X)},\mu_1^*(X)\right]\right| < \frac{\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]}$

or if,

b) $E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]$ *and* $cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]$ *are either both positive or both negative, and*

$$\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1^*(X)\right]\right| < \left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+\frac{cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}\right|$$

then,

$$\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)(\mu_1(X)-\mu_1^*(X))\right]\right| < \left|\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}-\mu_1\right|.$$

For a comparison between $\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$ and $\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$ the numerical example again shows that the necessary conditions of Theorem 14 and the sufficient condition in Theorem 15 a) are satisfied.

5. Simulation study

In order to investigate the asymptotic biases and also the finite sample performance of $\hat{\Delta}^*_{\mathrm{IPW}_1}$, $\hat{\Delta}^*_{\mathrm{IPW}_2}$ and $\hat{\Delta}^*_{\mathrm{DR}}$ under model misspecification of both the PS and the OR models we perform a simulation study with three different designs A, B and C. The first part of the simulations evaluate the finite sample performances of the estimators and consist of 1000 replications of sample sizes 500, 1000 and 5000. We generate covariates $X_1 \sim$ Uniform(1,4), $X_2 \sim$ Poisson(3) and $X_3 \sim$ Bernoulli(0.4). We use generalized linear models to generate a binary treatment T and potential outcomes $Y(t)$, $t = 0, 1$ with second order terms of X_1 and X_2 in both the PS and OR models. The PS-distributions for the treated and controls are bounded away from zero and 1 under the true models and under the model misspecifications. The PS and OR models (for the DR estimator) are stepwise misspecified. We have three designs where:

A: a quadratic term ${X_1}^2$ is omitted in the PS and OR models;
B: two quadratic terms, ${X_1}^2$ and ${X_2}^2$, are omitted in the PS and OR models;
C: two quadratic terms are omitted and the both the OR and PS link functions are misspecified.

The glm family and link functions together with the true parameter values are given in Table 1 which also contains the details for the misspecified models. The simulation is performed with the statistical software R (R Core Team 2015).

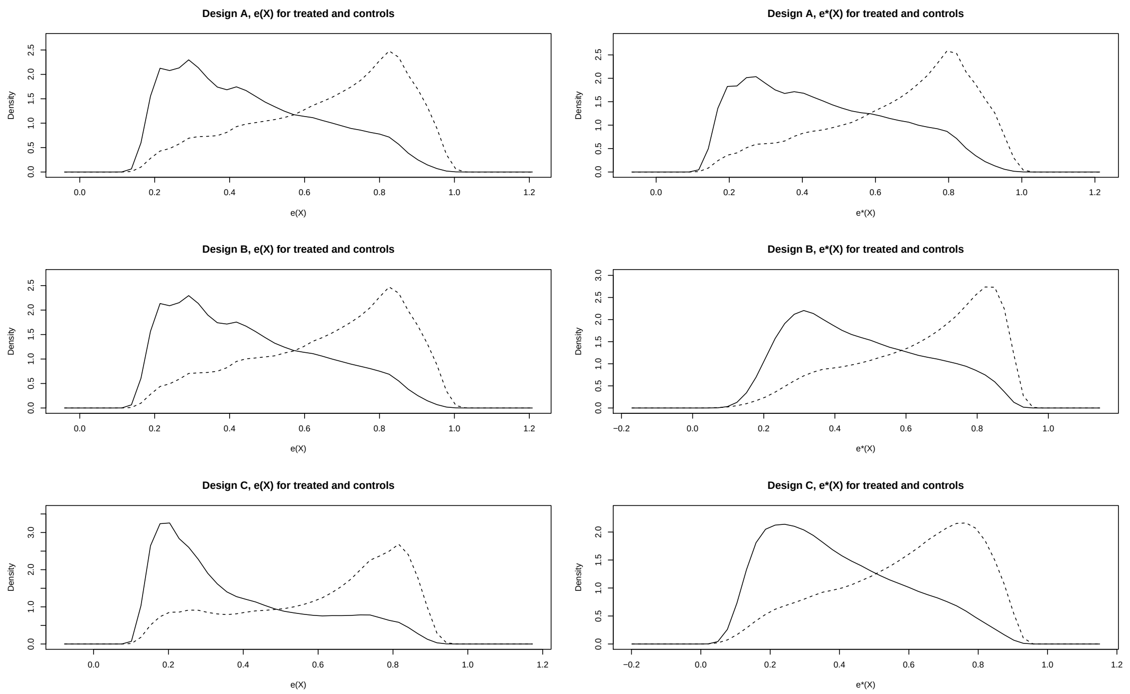
In Table 2 we give the simulation bias, standard error and MSE of the three estimators. When using the true models, i.e, when studying the estimators $\hat{\Delta}_{\mathrm{IPW}_1}$, $\hat{\Delta}_{\mathrm{IPW}_2}$ and $\hat{\Delta}_{\mathrm{DR}}$ the bias is small and decreases when the sample size increases and the standard errors follow the expected order with the smallest for $\hat{\Delta}_{\mathrm{DR}}$ followed by $\hat{\Delta}_{\mathrm{IPW}_2}$ and $\hat{\Delta}_{\mathrm{IPW}_1}$ (Lunceford and Davidian 2004). Under misspecification the bias does not decrease with the sample size but gets closer to the asymptotic biases, see Table 3. Under misspecification the standard errors follow the same pattern as under the true models. The bias of $\hat{\Delta}^*_{\mathrm{IPW}_1}$ is the largest whereas the $\hat{\Delta}^*_{\mathrm{IPW}_2}$ and $\hat{\Delta}^*_{\mathrm{DR}}$ are similar. The MSE of $\hat{\Delta}^*_{\mathrm{DR}}$ is the smallest, however for $n = 5000$, $\hat{\Delta}^*_{\mathrm{IPW}_2}$ and $\hat{\Delta}^*_{\mathrm{DR}}$ are very similiar.

In Table 3 we give numerical approximations for $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_1})$, $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_2})$ and $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{DR}})$ using a sample size of $n = 1,000,000$. We also show the same approximations for $\mathrm{Bias}_t(\hat{\Delta}^*_{\mathrm{IPW}_1})$, $\mathrm{Bias}_t(\hat{\Delta}^*_{\mathrm{IPW}_2})$ and $\mathrm{Bias}_t(\hat{\Delta}^*_{\mathrm{DR}})$, $t = 0, 1$. Here, we see that the total bias is smallest for $\hat{\Delta}^*_{\mathrm{IPW}_2}$ in Design A but smaller for $\hat{\Delta}^*_{\mathrm{DR}}$ in Design B and C. The absolute value of the biases in the two parts are smallest for $\hat{\Delta}^*_{\mathrm{DR}}$ in all designs. We also give the expectations and covariances that are used for the necessary and sufficient conditions in Theorems 11-15. We immediately see that the necessary condition for the absolute values of $\mathrm{Bias}_t(\hat{\Delta}^*_{\mathrm{DR}})$ to be smaller than the absolute value of $\mathrm{Bias}_t(\hat{\Delta}^*_{\mathrm{IPW}_1})$ holds for both $t = 0, 1$. The means (μ_0, μ_1) are close to the means under model misspecification (μ_0^*, μ_1^*) which is an assumption needed in order to evaluate the sufficient conditions in Theorems 13 a) and 15 a). By inspecting the covariances and the additional critera of the Theorems

13 a) and 15 a) in Table 3 we can see that the resulting inequalities of the theorems are in line with the results for the asymptotic biases.

Since we have that $E\left[(e(X)/e^{*}(X)-1)\,\mu_1(X)\right]$ and $cov\left[e(X)/e^{*}(X),\mu_1(X)\right]$ and $E\left[((1-e(X))/(1-e^{*}(X))-1)\,\mu_0(X)\right]$ and $cov\left[(1-e(X))(1-e^{*}(X)),\mu_0(X)\right]$ do not have the same sign in Design A and C the sufficient conditions in Theorems 13 b) and 15 b) cannot be applied.

Figure 1. Density plots of the propensity score distributions, $\hat{e}(X)$ and $\hat{e}^{*}(X)$ for treated and controls for Design A (top), B (middle), and C (bottom).



6. Discussion

In this paper we investigate biases of two IPW estimators and a DR estimator under model misspecification. For this purpose, we use a generic probability limit, under misspecification of the PS and OR models, which exists under general conditions. Since the propensity score enters the estimator in different ways for the IPW estimators under study the consequences of the model misspecification are not the same. The bias of the IPW estimators depend on the covariance between the PS-model error and the conditional outcome in different ways and the resulting bias can be in opposite directions. Comparing the bias of the DR estimator with a simple IPW estimator the necessary condition for the DR estimator to have a smaller bias is that the expectation of the outcome model under misspecification is less than twice the true conditional outcome, where the expectations also includes a scaling with the PS-model error. For the comparison with the normalized

IPW estimator the (PS-error scaled) misspecified outcome involves an interval defined by the true conditional outcome adding and subtracting the absolute value of the covariance between the PS-model error and the conditional outcome.

To our knowledge, there are only simulation studies comparing DR-estimators with other estimators (Kang and Schafer 2007; Waernbaum 2012) under the assumption that all models are misspecified. In this paper we study the same problem with an analytical approach although the comparisons are made between a DR estimator and IPW estimators.

References

Bang, H. and J. M. Robins (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics 61*(4), 962–973.

Boos, D. D. and L. Stefanski (2013). M-estimation (estimating equations). In *Essential Statistical Inference*, Volume 120, pp. 297–337. Springer New York.

Busso, M., J. DiNardo, et al. (2014). New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics 96*(5), 885–897.

Cao, W., A. A. Tsiatis, et al. (2009). Improving efficiency and robustness of the doubly robust estimator for a population mean with incomplete data. *Biometrika 96*(3), 723–734.

Hernán, M. A., B. Brumback, et al. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of hiv-positive men. *Epidemiology 11*(5), 561–570.

Hernán, M. A. and J. M. Robins (2006). Estimating causal effects from epidemiological data. *Journal of Epidemiology and Community Health 60*(7), 578–586.

Hirano, K. and G. W. Imbens (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology 2*(3), 259–278.

Hirano, K., G. W. Imbens, et al. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica 71*(4), 1161–1189.

Imai, K. and M. Ratkovic (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76*(1), 243–263.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

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

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

Kwon, T., I. G. Jeong, et al. (2015). Adjuvant chemotherapy after radical cystectomy for bladder cancer: a comparative study using inverse-probability-of-treatment weighting. *Journal of Cancer Research and Clinical Oncology 141*(1), 169–176.

Lunceford, J. K. and M. Davidian (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine 23*(19), 2937–2960.

Petersen, M. L., K. E. Porter, S. Gruber, Y. Wang, and M. J. van der Laan (2010). Diagnosing and responding to violations in the positivity assumption. *Statistical Methods in Medical Research 21*(1), 31–54.

R Core Team (2015). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Robins, J., M. Sued, et al. (2007). Comment: Performance of double-robust estimators when "inverse probability" weights are highly variable. *Statistical Science 22*(4), 544–559.

Robins, J. M. and A. Rotnitzky (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association 89*(427), 846–866.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*(5), 688–701.

Scharfstein, D. O., A. Rotnitzky, et al. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association 94*(448), 1096–1120.

Seaman, S. R. and I. R. White (2013). Review of inverse probability weighting for dealing with missing data. *Statistical Methods in Medical Research 22*(3), 278–295.

Tan, Z. (2007). Comment: Understanding OR, PS and DR. *Statistical Science 22*(4), 560–568.

Tsiatis, A. A. (2007). *Semiparametric theory and missing data.* Springer Science & Business Media.

Tsiatis, A. A. and M. Davidian (2007). Comment: Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science 22*(4), 569–573.

Vansteelandt, S., J. Carpenter, et al. (2010). Analysis of incomplete data using inverse probability weighting and doubly robust estimators. *Methodology 6*(1), 37–48.

Vermeulen, K. and S. Vansteelandt (2015). Bias-reduced doubly robust estimation. *Journal of the American Statistical Association 110*(511), 1024–1036.

Waernbaum, I. (2012). Model misspecification and robustness in causal inference: comparing matching with doubly robust estimation. *Statistics in Medicine 31*(15), 1572–1581.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society 50*(1), 1–25.

Williamson, E., A. Forbes, and I. White (2014). Variance reduction in randomised trials by inverse probability weighting using the propensity score. *Statistics in Medicine 33*(1), 721–737.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Yao, L., Z. Sun, et al. (2010). Estimation of average treatment effects based on parametric propensity score model. *Journal of Statistical Planning and Inference 140*(3), 806–816.

Appendix A

A.1. **Regularity conditions for applying a weak law of large numbers for averages of functions with estimated parameters.** The convergence in probability of $\hat{\Delta}^*_{\mathrm{IPW}_1}$, $\hat{\Delta}^*_{\mathrm{IPW}_2}$ and $\hat{\Delta}^*_{\mathrm{DR}}$ to their corresponding expectations would follow directly from a WLLN for an iid sample of (T_i, X_i, Y_i) except for the estimated parameters $\hat{\beta}^*$ in $\hat{e}^*(X_i)$ and and $\hat{\alpha}^*_t$ in $\hat{\mu}^*_t(X_i)$, $t = 0, 1$. To justify the biases in Section 4 consider a general representation of a function $g\left[T, Y, X, \hat{\theta}\right]$ where $\hat{\theta} \xrightarrow{p} \theta_0$ and

$$\frac{1}{n}\sum g\left[X_i, T_i, Y_i, \hat{\theta}\right] \xrightarrow{p} E\left[g(T, X, Y, \theta_0)\right] \tag{13}$$

The $\hat{\theta}$ in (13) corresponds to $\hat{\beta}^*$ for $\hat{\Delta}^*_{\mathrm{IPW}_1}$, $\hat{\Delta}^*_{\mathrm{IPW}_2}$ and $(\hat{\alpha}^*, \hat{\beta}^*)$ for $\hat{\Delta}^*_{\mathrm{DR}}$ and under Assumptions 5 and 7 the consistency of $\hat{\theta}$ is ensured. Regularity conditions for the function g can be given see e.g., Boos and Stefanski (2013, Theorem 7.3) who show that (13) holds for differentiable functions with bounded derivatives (wrt θ). The regularity conditions for $g\left[X_i, T_i, Y_i, \hat{\theta}\right]$, for the three estimators, imply conditions on the models $e^*(X, \beta^*)$ and $\mu^*_t(X, \alpha^*_t)$ such that the regularity condition for g is satisfied. Under (13) we can insert the limiting values β^* and α^* and their corresponding $e^*(X)$ and $\mu^*_t(X)$, $t = 0, 1$ when taking a WLLN.

A.2. **Biases.** Under the regularity conditions and Assumptions 5-6 (IPW) and 5-7 (DR) we derive the bias in Theorem 8 below. For $\hat{\Delta}^*_{\mathrm{IPW}_1}$:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{T_iY_i}{\hat{e}^*(X_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-T_i)Y_i}{1-\hat{e}^*(X_i)} \xrightarrow{p} E\left[\frac{TY}{e^*(X)}\right] - E\left[\frac{(1-T)\,Y}{(1-e^*(X))}\right],$$

and

$$\begin{aligned} E\left[\frac{TY}{e^*(X)}\right] - E\left[\frac{(1-T)\,Y}{(1-e^*(X))}\right] &= E\left[E\left[\frac{TY}{e^*(X)}\middle| X\right]\right] - E\left[E\left[\frac{(1-T)\,Y}{[1-e^*(X)]}\middle| X\right]\right] \\ &= E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - E\left[\frac{(1-e(X))}{(1-e^*(X))}\mu_0(X)\right]. \end{aligned}$$

subtracting with Δ gives

$$\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_1}) = E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right] - (\mu_1 - \mu_0).$$

The biases of Theorems 9 and 10 are derived similarly.

A.3. **Comparisons.** To study the consequences of model misspecification for the estimators we compare each difference involving $\mu_1(X)$ and $\mu_0(X)$ separately. For example we study $\mathrm{Bias}(\hat{\Delta}^*_{\mathrm{IPW}_1})$

$$\hat{\Delta}^*_{\mathrm{IPW}_1} - \Delta \xrightarrow{p} E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right] - (\mu_1 - \mu_0)$$
$$= E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - \mu_1 + \mu_0 - E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X))\right]$$

and Bias($\hat{\Delta}^*_{\mathrm{IPW}_2}$)

$$\hat{\Delta}^*_{\mathrm{IPW}_2} - \Delta \xrightarrow{p} \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \frac{E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right]}{E\left[\frac{1-e(X)}{1-e^*(X)}\right]} - (\mu_1 - \mu_0)$$
$$= \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1 + \mu_0 - \frac{E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right]}{E\left[\frac{1-e(X)}{1-e^*(X)}\right]}.$$

The inequalities concerning the biases are made with respect to the absolute values for two parts separately, i.e, for Theorem 11

$$\left|E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right] - \mu_1\right| < \left|\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1\right|. \tag{14}$$

Since $|-a| = |a|$ the same comparison for the second part are

$$\left|E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right] - \mu_0\right| < \left|\frac{E\left[\frac{1-e(X)}{1-e^*(X)}\mu_0(X)\right]}{E\left[\frac{1-e(X)}{1-e^*(X)}\right]} - \mu_0\right|, \tag{15}$$

and the conditions derived for the first part of the biases, (10) and (11), can be directly applied to (15) replacing $e(X)/e^*(X)$, with $(1-e(X))/(1-e^*(X))$ and $\mu_1(X)$ with $\mu_0(X)$. Similarly, for the second part of the bias in Theorem 10 we have that

$$\left|E\left[\frac{(1-e(X))(1-e^*(X))\,(\mu_0(X) - \mu_0^*(X))}{(1-e^*(X))}\right]\right| = \left|E\left[\frac{(e(X) - e^*(X))\,(\mu_0(X) - \mu_0^*(X))}{(1-e^*(X))}\right]\right|$$

and the conditions derived for the first part of the bias, (12), can be directly applied to (15) additionally replacing $\mu_1^*(X)$ with $\mu_0^*(X)$.

Proof of Theorem 12. 1. Assuming that $E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] > 0$:

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right| < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right],$$

Here, we have

$$-E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right],$$

and

$$0 < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right] < 2 \cdot E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]. \tag{16}$$

2. Assuming that $E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] < 0$:

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right| < -E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right],$$

Here, we have

$$E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right] < -E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right],$$

(17)

$$2 \cdot E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right] < 0$$

and by (16) and (17)

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right]\right| \leq 2 \cdot \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|.$$

□

Proof of Theorem 13. a) We have that

$$\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] - cov\left[\frac{e(X)}{e^*(X)}, \mu_1^*(X)\right]\right| = \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right|.$$

by $\mu_1^* = \mu_1$. Also,

$$\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\frac{e(X)}{e^*(X)} - 1\right]E\left[\mu_1(X)\right]\right| = \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|.$$

Further, assuming that

$$\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] - cov\left[\frac{e(X)}{e^*(X)}, \mu_1^*(X)\right]\right| < \left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] + E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1\right|,$$

it follows that

$$\left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right| < \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]\right|.$$

c) see proof of Theorem 12.

□

Proof of Theorem 14. 1. Assuming that:

$$cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] > 0 \iff \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1 > 0$$

Here, we have

$$-\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} + \mu_1 < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right] < \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1,$$

$$-cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] < E\left[\frac{e(X)}{e^*(X)}\right] E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right] < cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right],$$

and

(18)

$$E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] - \frac{cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] + \frac{cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}$$

2. Assuming that:

$$cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] < 0 \iff \frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1 < 0,$$

then it follows that

$$\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} - \mu_1 < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right] < -\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} + \mu_1$$

$$cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] < E\left[\frac{e(X)}{e^*(X)}\right] E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right] < -cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right],$$

and

(19)

$$E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] + \frac{cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]} < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] - \frac{cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]},$$

hence by (18) and (19) we conclude that

$$E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] - \frac{\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]} < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right] < E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right] + \frac{\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]},$$

□

Proof of Theorem 15. For a) assuming $\mu_1^* = \mu_1$ implies

$$\left|cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right] - cov\left[\frac{e(X)}{e^*(X)}, \mu_1^*(X)\right]\right| = \left|E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)(\mu_1(X) - \mu_1^*(X))\right]\right|.$$

Also,

$$\frac{\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]}=\left|\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}-\mu_1\right|,$$

and by

$$\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]-cov\left[\frac{e(X)}{e^*(X)},\mu_1^*(X)\right]\right|<\frac{\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|}{E\left[\frac{e(X)}{e^*(X)}\right]},$$

it follows that

$$\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)(\mu_1(X)-\mu_1^*(X))\right]\right|<\left|\frac{E\left[\frac{e(X)}{e^*(X)}\mu_1(X)\right]}{E\left[\frac{e(X)}{e^*(X)}\right]}-\mu_1\right|.$$

For b) we use that if $E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]$ and $cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]$ are both positive then,

$$E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|=\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|$$

and

$$-\left|E\left\{\left[\frac{e(X)}{e^*(X)}-1\right]\mu_1(X)\right\}+cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|<E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|.$$

If $E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]$ and $cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]$ are both negative, then we have

$$E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|<\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|,$$

and

$$-\left|E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|=E\left[\left(\frac{e(X)}{e^*(X)}-1\right)\mu_1(X)\right]+\left|cov\left[\frac{e(X)}{e^*(X)},\mu_1(X)\right]\right|,$$

and the necessary condition from Theorem 14 follows. □

Table 1. Simulation Designs A, B and C

	TRUE MODEL		MISSPECIFIED MODEL	
Models	Class	Linear predictor and parameter values	Class	Linear predictor
		$\beta = (-1, 0.6, 0.1, 0.9, 0.1, 0.7), \alpha_0 = (3, 0.5, 0.2, 0.5, 0.2, 0.2)$ $\alpha_1 = (4, 1.1, 0.1, 0.5, 0.3, 0.2)$		
Design A				
PS	Binomial, logit	$X_1, X_2, X_1^2, X_2^2, X_3$	Binomial, logit	X_1, X_2, X_2^2, X_3
OR	Gaussian, identity	$X_1, X_2, X_1^2, X_2^2, X_3$	Gaussian, identity	X_1, X_2, X_2^2, X_3
Design B				
PS	Binomial, logit	$X_1, X_2, X_1^2, X_2^2, X_3$	Binomial, logit	X_1, X_2, X_3
OR	Gaussian, identity	$X_1, X_2, X_1^2, X_2^2, X_3$	Gaussian, identity	X_1, X_2, X_3
Design C				
PS	Binomial, cauchit	$X_1, X_2, X_1^2, X_2^2, X_3$	Binomial, logit	X_1, X_2, X_3
OR	Gamma, identity	$X_1, X_2, X_1^2, X_2^2, X_3$	Gaussian, identity	X_1, X_2, X_3

Table 2. Simulation results

			ESTIMATORS								
			$\hat{\Delta}^*_{\mathrm{IPW}_1}$			$\hat{\Delta}^*_{\mathrm{IPW}_2}$			$\hat{\Delta}^*_{\mathrm{DR}}$		
n	Models	Design	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
500	True		0.018	0.368	0.136	0.013	0.138	0.019	$<$0.001	0.11	0.012
	False	A	0.118	0.390	0.166	0.019	0.136	0.019	0.021	0.119	0.015
	False	B	0.262	0.324	0.174	0.033	0.126	0.017	0.024	0.110	0.013
	False	C	0.229	0.349	0.174	-0.061	0.128	0.020	0.037	0.115	0.015
1000	True		0.010	0.254	0.065	0.002	0.098	0.010	$<$ 0.001	0.075	0.006
	False	A	0.112	0.270	0.085	0.008	0.097	0.009	0.018	0.079	0.007
	False	B	0.260	0.214	0.113	0.034	0.088	0.009	0.025	0.078	0.007
	False	C	0.243	0.222	0.108	-0.056	0.092	0.012	0.043	0.081	0.008
5000	True		0.007	0.107	0.011	0.005	0.044	0.002	0.003	0.035	0.001
	False	A	0.112	0.116	0.026	0.013	0.044	0.002	0.025	0.037	0.002
	False	B	0.260	0.090	0.076	0.033	0.039	0.003	0.025	0.036	0.002
	False	C	0.226	0.096	0.060	-0.057	0.040	0.005	0.040	0.036	0.003

TABLE 3. Asymptotic approximations from Designs A, B and C.

Parameter	Design A	Design B	Design C
μ_1	11.127	11.127	12.130
μ_1^*	11.092	11.098	12.097
μ_0	8.628	8.628	9.633
μ_0^*	8.578	8.564	9.582
$\text{Bias}(\hat{\Delta}^*_{\text{IPW}_1})$	0.096	0.264	0.213
$\text{Bias}(\hat{\Delta}^*_{\text{IPW}_2})$	0.007	0.033	-0.057
$\text{Bias}(\hat{\Delta}^*_{\text{DR}})$	0.017	0.025	0.037
$\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_1})$	0.024	0.128	0.130
$\text{Bias}_1(\hat{\Delta}^*_{\text{IPW}_2})$	-0.028	0.013	-0.089
$\text{Bias}_1(\hat{\Delta}^*_{\text{DR}})$	0.009	0.009	0.029
$E\left[\frac{e(X)}{e^*(X)}\right]$	1.005	1.010	1.019
$cov\left[\frac{e(X)}{e^*(X)}, \mu_1(X)\right]$	-0.029	0.015	-0.095
$cov\left[\frac{e(X)}{e^*(X)}, \mu_1^*(X)\right]$	-0.040	0.006	-0.121
$E\left[\frac{e(X)}{e^*(X)} - 1\right]\mu_1$	0.060	0.113	0.224
$E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1(X)\right]$	0.030	0.128	0.130
$E\left[\left(\frac{e(X)}{e^*(X)} - 1\right)\mu_1^*(X)\right]$	0.019	0.119	0.102
$\text{Bias}_2(\hat{\Delta}^*_{\text{IPW}_1})$	0.076	0.137	0.083
$\text{Bias}_2(\hat{\Delta}^*_{\text{IPW}_2})$	0.039	0.022	0.031
$\text{Bias}_2(\hat{\Delta}^*_{\text{DR}})$	0.011	0.018	0.008
$E\left[\frac{1-e(X)}{1-e^*(X)}\right]$	0.995	0.987	0.995
$cov\left[\frac{1-e(X)}{1-e^*(X)}, \mu_0(X)\right]$	-0.037	-0.017	-0.033
$cov\left[\frac{1-e(X)}{1-e^*(X)}, \mu_0^*(X)\right]$	-0.026	-0.004	-0.024
$E\left[\frac{1-e(X)}{1-e^*(X)} - 1\right]\mu_0$	-0.038	0.111	-0.052
$E\left[\left(\frac{1-e(X)}{1-e^*(X)} - 1\right)\mu_0(X)\right]$	-0.074	-0.128	-0.085
$E\left[\left(\frac{1-e(X)}{1-e^*(X)} - 1\right)\mu_0^*(X)\right]$	-0.065	-0.114	-0.075