Journal of Statistical Research *ISSN 0256 - 422 X*
2018, Vol. 52, No. 1, pp. 91-113

VARIANTS OF DOUBLE ROBUST ESTIMATORS FOR TWO-STAGE DYNAMIC TREATMENT REGIMES

ANDREW S. TOPP

AbbVie Inc. North Chicago, IL, U.S.A.

Email: andrew.topp@abbvie.com

GEOFFREY S. JOHNSON

GlaxoSmithKline, Collegeville, PA, U.S.A.

Email: geoffrey.s.johnson@gsk.com

ABDUS S. WAHED*

Department of Biostatistics, University of Pittsburgh Graduate School of Public Health Pittsburgh, PA, U.S.A.

Email: WahedA@edc.pitt.edu

SUMMARY

Certain conditions and illnesses may necessitate multiple stages of treatment and thus require unique study designs to compare the efficacy of these interventions. Such studies are characterized by two or more stages of treatment punctuated by decision points where intermediate outcomes inform the choice for the next stage of treatment. The algorithm that dictates what treatments to take based on intermediate outcomes is referred to as a dynamic regime. This paper describes an efficient method of building double robust estimators of the treatment effect of different regimes. A double robust estimator utilizes both modeling of the outcome and weighting based on the modeled probability of receiving treatment in such a way that the resulting estimator is a consistent estimate of the desired population parameter under the condition that at least one of those models is correct. This new and more efficient double robust estimator is compared to another double robust estimator as well as classical regression and inverse probability weighted estimators. The methods are applied to estimate the regime effects in the STAR*D anti-depression treatment trial.

Keywords and phrases: Double robustness; Dynamic treatment regime; Inverse probability weighting

* Corresponding author

1 Introduction

When comparing two or more treatments in observational or randomized studies, for a given patient, investigators can only observe the outcome corresponding to the treatment received by the patient. Thus, investigators are unable to observe the potential outcomes that could have been observed had a patient received other treatments. These potential outcomes a patient could have experienced are often referred to as counterfactual outcomes (Holland, 1986).

While counterfactual variables are not always observable, they provide avenues to construct valid estimates and to assess treatment effects from randomized trials or observational studies. Various methods have been proposed that produce unbiased and consistent estimates of treatment effects based on counterfactuals. The standard and most straightforward method is to directly model the outcome given receiving each treatment. However, modeling an outcome may not be straightforward because of the uncertainty of the relationship between the observed covariates and the outcome. In the case of non-randomized studies, there may be unobserved confounders important to the causal relationship between treatment and outcome. An alternate approach is called inverse probability of treatment weighting. In this approach each observation is weighted with the inverse of the probability of receiving the treatment they received. So a patient who received treatment A_1 would have a weight of $1/Pr(ReceivingA_1)$. With this method, a patient counts for herself along with $1/Pr(ReceivingA_1) - 1$ other patients that have similar covariate values but for whom we did not see the counterfactual outcome given treatment A_1. In classical randomized trials, these probabilities are known and hence it is straightforward to apply inverse probability of treatment weighting. In observational studies, however, treatment assignment probabilities are not known, and needs to be estimated from the data. If we model the probability of receiving treatment based on the observed covariates the analysis may be flawed in the same way as when modeling the outcome incorrectly in the previous method. In this case, the accuracy of the estimate will depend on the relationship between the probability of receiving treatment and the included covariates. Incorrect specification of this relationship could lead to biased and/or inefficient inference. However, there are methods requiring the specification of both outcome and treatment models, called double-robust methods, that allow for unbiased or consistent estimates. This continues to be true if at least one of the models is correctly specified. Double robust estimation will be the focus of this paper.

There has been a wealth of literature published on the subject of double robust estimation since its introduction by Robins (2000). Carpenter et al. (2006) compares this method to the more common approach of multiple imputation. Both Bang and Robins (2005) and Tsiatis et al. (2011) utilize double robust estimators in a longitudinal setting and attempt to improve to this method's efficiency. Kang and Schafer (2007) compare various double robust, as well as some non-double robust methods in different missing data scenarios. Rotnitzky et al. (2012) discuss a double robust estimator derived by solving outcome regression estimating equations. Bai et al. (2013) examine a double robust estimator in the survival data setting, with particular emphasis on stratified sampling schemes.

Cao et al. (2009) in particular are concerned with the efficient estimation of the outcome model coefficients in cross-sectional settings. They observe that even if one correctly specified the treatment model with incorrect specification of the outcome model, the results, while still yielding consistent estimates, will be inefficient. They propose an estimating equation for these coefficients that

results in more efficient double robust estimates.

Double robust estimators discussed above were mostly derived or implemented for settings where one treatment is compared to another in cross-sectional or longitudinal settings. But chronic diseases demand frequent modification of treatments based on individual responses to the treatment and/or health conditions. Therefore, instead of learning which treatment is immediately better, a physician may want to know what course of treatments is best overall. These courses of treatments are called dynamic treatment regimes (DTRs). They are rules of treatment choice based on intermediate responses and patient characteristics. For example, a dynamic treatment regime in depression treatment could be 'treat a patient suffering from clinical depression with a selective serotonin reuptake inhibitor (SSRI), if they respond keep them on it, and if not switch them to a drug like bupropion (a non-SSRI treatment for clinical depression)'.

Comparing dynamic treatment regimes using counterfactual outcomes is generally more complicated than comparing a single set of treatments. For the latter it is enough to consider counterfactuals related to a group of treatments, but we need to think about all possible combinations of initial and follow-up treatments a patient could experience in a dynamic treatment regime setting. Dynamic treatment regimes and related analytical techniques have grown in prominence in the past two decades. Murphy (2003) discusses the use of backwards induction and Q-learning as a technique for estimating regime outcomes. Bembom and van der Laan (2007) explore both inverse probability weighting and g-computation in two stage dynamic treatment regime clinical trials. Lunceford et al. (2002) discuss an inverse probability of treatment weighted estimator with a correctional constant chosen to minimize variance across a class of estimators they introduce. Zhang et al. (2013) propose a double robust estimator in a dynamic treatment regime setting whose efficiency improves through careful choice of an augmentation term.

Many well known, large studies aim to compare dynamic treatment regimes. The STAR*D trial (Warden et al., 2007), which motivated this research, is a multistage clinical trial examining the effects of depression treatments for patients who do not respond to SSRIs. The COG study A3891 (Matthay et al., 1999) examines the efficacy of chemotherapy followed by bone marrow transplantation versus a second round of chemotherapy in children with high-risk neuroblastoma. The REVAMP study (Trivedi et al., 2008) is another multistage study concerned with analyzing regimes of treatments for chronic depression.

In this paper we examine double robust estimators in a dynamic treatment regime setting. In particular, we propose a double robust estimator that is more efficient then the existing estimators. We present a comparative simulation study, and demonstrate the methods by applying them to analyze STAR*D study data.

2 Framework

2.1 Estimating Causal Treatment Effect

As mentioned in the section above when comparing treatments within a study, investigators can only observe the outcome corresponding to the treatment the patient was assigned. So, the potential

outcome the patient would have experienced had the patient been assigned to the other treatment is missing. One common approach to dealing with missing data is to replace the missing values with an approximation based on the observed data. In regression-based imputation, we construct a model of the outcome regressed on treatment and any relevant covariates, which allows us to estimate the conditional expectation of the outcome given the treatment the patient did not receive using the resulting model. An unbiased, or at least consistent estimate of the mean outcome conditional on receiving a particular treatment, $\mu = E[Y_i|A]$, is then constructed by these conditional expectations. More specifically, $\hat{\mu}_{OR} = (1/n)\sum_{i=1}^{n} m(X_i, \hat{\beta})$, where $m(X_i, \beta)$ is the postulated ordinary least squares regression model for $E[Y_i|X_i]$, n is the number of subjects in the sample, and $\hat{\beta}$ is a consistent estimator of β, the parameter for the outcome model.

An alternate approach would be to impute the missing outcomes using the outcomes of other, similar patients. Assume a patient received treatment A with the probability of receiving it $\pi^{(A)}(X_i, \gamma)$, where γ are coefficients that describe the relationship between the covariates, X_i, and probability of receiving treatment. Then, on average, there are $[1/\pi^{(A)}(X_i, \gamma) - 1]$ patients with covariates X_i received a different treatment. Therefore, by weighting this patient by $1/\pi^{(A)}(X_i, \gamma)$ we can use their outcome to account for themselves as well as those $[1/\pi^{(A)}(X_i, \gamma) - 1]$ other patients. Using this method to estimate the mean outcome conditional on receiving A would result in the Inverse Probability of Treatment Weighting estimator (Horvitz and Thompson, 1952).

$$\hat{\mu}_{IPTW} = \frac{1}{n}\sum_{i=1}^{n} \frac{Z_i^{(A)}}{\pi^{(A)}(X_i, \hat{\gamma})} Y_i, \tag{2.1}$$

where $Z_i^{(A)}$ is the indicator for receiving treatment A, $Z_i^{(A)} = 1$, if the patient receives A, 0, otherwise, and Y_i is the outcome for the i^{th} subject, $i = 1, \ldots, n$. In the above equation we have also replaced γ by $\hat{\gamma}$ to indicate that when $\pi^{(A)}$ is unknown, it is estimated from the observed data.

Both of these estimating approaches provide a consistent estimator of the population mean μ given the correct specification of the models involved, namely, outcome and treatment. Misspecification of these models may result in biased and inefficient estimates. However, a third approach, double robust estimation, allows for consistent estimation as long as at least one of these models is correctly specified. This double robust estimator is given by

$$\hat{\mu}_{DR} = \frac{1}{n}\sum_{i=1}^{n} \left\{ \frac{Z_i^{(A)}}{\pi^{(A)}(X_i, \hat{\gamma})} Y_i - \frac{Z_i^{(A)} - \pi^{(A)}(X_i, \hat{\gamma})}{\pi^{(A)}(X_i, \hat{\gamma})} m(X_i, \hat{\beta}) \right\}. \tag{2.2}$$

For a formal argument for why this estimator is consistent when at least one of the two models is correctly specified, we refer readers to Robins (2000).

2.2 Inference for Dynamic Treatment Regimes

Now consider the setting of two-stage dynamic treatment regimes where, at the second stage, nonresponders receive subsequent follow-up treatments and responders continue to follow-up. Patients

will receive up to two levels of treatment. For simplicity, suppose there are only two first stage treatment options, A_1 and A_2, and two second stage treatment options, B_1 and B_2, for non-responders to the first stage treatments A_1 or A_2. We denote intermediate response with $R_i = 1$ denoting the i^{th} patient response and $R_i = 0$ for their non-response after the first stage of treatment. This setting allows four dynamic treatment regimes, namely, $d(A_j, B_k)$, $j, k = 1, 2$, where, for example, in regime $d(A_1, B_1)$, a patient initially receives treatment A_1 and if non-responsive is switched to treatment B_1. A patient who receives A_1, responds, and stays on that treatment would be considered to be on this regime, and so would someone that does not respond to A_1 and is switched to B_1. This creates an interesting problem when trying to estimate the mean regime outcome in even a randomized controlled trial. If there were another second stage treatment that a non-responder could switch to, B_2, then for patients who did not respond to A_1 and were switched to B_2, we do not observe the outcome under regime $d(A_1, B_1)$. However, they were part of that regime up until the second stage of treatment. This is equivalent to outcome data missing for these patients under regime $d(A_1, B_1)$ and hence when estimating the mean outcome under $d(A_1, B_1)$, we need to account for these missing data. If we simply take the average outcome of all patients who can be described as following regime $d(A_1, B_1)$ then we may get a biased estimate. In this average responders to A_1 will count for more than non-responders because some portion of non-responders to A_1, those assigned to B_2, have their outcome missing for our regime of interest.

Let us cast our problem in terms of counterfactual outcomes. Suppose $Y[d(A_j, B_k)]$ denotes the outcome had the patient been treated with the regime $d(A_j, B_k)$. In terms of this, the goal is to estimate $\mu[d(A_j, B_k)] = E\{Y[d(A_j, B_k)]\}$, the mean of a population who were treated with $d(A_j, B_k)$. However, $Y_i[d(A_j, B_k)]$, for all j,k=1,2, is not observed for every individual. Instead we only observe

$$\left\{X_i, Z_{ji}^{(A)}, R_i, (1 - R_i)Z_{ki}^{(B)}, Y_i; i = 1, \ldots, n; j, k = 1, 2\right\}, \tag{2.3}$$

where X_i is the patient's observed covariates, $Z_{ji}^{(A)}$ is the first stage treatment assignment indicator for treatment A_j, R_i is the response indicator for the first stage treatment, $Z_{ki}^{(B)}$ is the second stage treatment assignment indicator for non-responders receiving treatment B_k, and Y_i is the outcome.

The primary assumption required for causal inference is consistency (Robbins et al., 2000), which in a dynamic treatment regime setting means that the outcome Y_i a patient experiences under a regime $d(A_j, B_k)$ is in fact the counterfactual outcome $Y_i[d(A_j, B_k)]$. In the case of two treatments at each stage this would be

$$\begin{aligned} Y_i = Z_{1i}^{(A)}\left[R_i + (1 - R_i) * Z_{1i}^{(B)}\right] * Y_i\left[d(A_1, B_1)\right] \\ + \left(1 - Z_{1i}^{(A)}\right)\left[R_i + (1 - R_i) * Z_{1i}^{(B)}\right] * Y_i\left[d(A_2, B_1)\right] \\ + Z_{1i}^{(A)}\left[R_i + (1 - R_i) * (1 - Z_{1i}^{(B)})\right] * Y_i\left[d(A_1, B_2)\right] \\ + \left(1 - Z_{1i}^{(A)}\right)\left[R_i + (1 - R_i) * (1 - Z_{1i}^{(B)})\right] * Y_i\left[d(A_2, B_2)\right]. \end{aligned} \tag{2.4}$$

Another assumption necessary for making causal inference is positivity. That is, that the probability of each treatment pathway being experienced by a patient is non-zero. Staying in our two treatment

per stage example this would mean

$$\begin{aligned} 1 > P(Z_{ji}^{(A)} = 1) > 0, j = 1, 2 \\ 1 > P(R_i = 1 | Z_{ji}^{(A)} = 1) > 0, j = 1, 2 \\ 1 > P(Z_{ki}^{(B)} = 1 | Z_{ji}^{(A)} = 1, R_i = 0) > 0, \forall j, k = 1, 2. \end{aligned} \tag{2.5}$$

The last of the three assumptions is sequential randomization. This assumption states that the probability of being assigned to a treatment at a decision point depends only on information observed up to that time. More explicitly,

$$\begin{aligned} P(Z_{ji}^{(A)} = 1) = P(Z_{ji}^{(A)} = 1 | X_i) \\ P(Z_{ki}^{(B)} = 1) = P(Z_{ki}^{(B)} = 1 | X_i, Z_{ji}^{(A)}). \end{aligned} \tag{2.6}$$

Bembom and van der Laan (2007) describe a method, known as g-computation, originally described by Robins (1986), that uses weighted averages, based on response probability, of conditional intermediate responses to achieve consistent estimates of mean regime outcomes. Specifically, the mean outcome under regime $d(A_j, B_k)$, $\mu[d(A_j, B_k)] = E[Y[d(A_j, B_k)]]$, is estimated as

$$\hat{\mu}_G[d(A_j, B_k)] = \frac{1}{n} \sum_{i=1}^{n} \left\{ R_i m_j(X_i, \hat{\beta}_j) \ + (1 - R_i) m_{jk}(X_i, \hat{\beta_{jk}}) \right\}, \tag{2.7}$$

where R_i is the i^{th} patient's response after the first stage of treatment, $R_i = 1$, if the responded to first stage treatment, 0 otherwise, X_i is the i^{th} patient's covariate information, X_i, $m_j(X_i, \beta_j)$ is the postulated outcome model for the conditional expectation of Y for responders given the covariates X_i, initial treatment A_j. Likewise, $m_{jk}(X_i, \beta_{jk})$ is the postulated outcome model for the conditional expectation of Y for non-responders given the covariates X_i, initial treatment A_j and second stage treatment B_k. We used hats to indicate that they are now estimated from the observed data.

Following the principle of inverse weighting described in Section 2.1, the inverse probability of treatment weighted estimator in a two stage dynamic treatment regime setting, as described by Wahed and Tsiatis (2004), becomes

$$\hat{\mu}_{IPTW}[d(A_j, B_k)] = \frac{1}{n} \sum_{i=1}^{n} \left\{ \frac{Z_{ji}^{(A)}}{\pi_j^{(A)}(X_i, \hat{\gamma_j})} \left[R_i + \frac{(1 - R_i) Z_{ki}^{(B)}}{\pi_{jk}^{(B)}(X_i, \hat{\gamma_{jk}})} \right] \right\} Y_i \ , \tag{2.8}$$

where $\pi_j^{(A)}(X_i, \hat{\gamma_j})$ is the modeled probability of receiving treatment A_j based on the covariates X_i, and $\pi_{jk}^{(B)}(X_i, \hat{\gamma_{jk}})$ is the modeled probability of receiving treatment B_k based on the covariates X_i given no response to $Z_{ji}^{(A)}$.

Lastly consider a locally efficient double robust estimator for a two-stage dynamic treatment

regime introduced by Wahed and Tsiatis (2004).

$$\hat{\mu}_{DR}[d(A_j,B_k)] = \sum_{i=1}^{n} Z_{ji}^{(A)} \left\{ \left[R_i + \frac{(1-R_i)Z_{ki}^{(B)}}{\pi_{jk}^{(B)}(X_i,\hat{\gamma_{jk}})} \right] Y_i \right. \\ \left. - \frac{Z_{ki}^{(B)} - \pi_{jk}^{(B)}(X_i,\hat{\gamma_{jk}})}{\pi_{jk}^{(B)}(X_i,\hat{\gamma_{jk}})}(1-R_i)m_{jk}(X_i,\hat{\beta}_{jk}) \right\} / \sum_{i=1}^{n} Z_{ji}^{(A)} \,. \quad (2.9)$$

This estimator is consistent provided at least one of the two models π_{jk} and m_{jk} are correctly specified. When both models are correctly specified, this estimator is most efficient, provided β_j and β_{jk} are consistently estimated. An extension of this estimator to include an inverse probability of treatment weighting component for first stage of treatment A_j this estimator results in

$$\hat{\mu}_{DR}[d(A_j,B_k)] = \frac{1}{n}\sum_{i=1}^{n} \left\{ \frac{Z_{ji}^{(A)}}{\hat{\pi}_j^{(A)}} \left[R_i + \frac{(1-R_i)Z_{ki}^{(B)}}{\hat{\pi}_{jk}^{(B)}} \right] Y_i - \frac{Z_{ji}^{(A)} - \hat{\pi}_j^{(A)}}{\hat{\pi}_j^{(A)}} R_i m_j(X_i,\hat{\beta}_j) \right. \\ \left. - \frac{Z_{ji}^{(A)} Z_{ki}^{(B)} - \hat{\pi}_j^{(A)}\hat{\pi}_{jk}^{(B)}}{\hat{\pi}_j^{(A)}\hat{\pi}_{jk}^{(B)}}(1-R_i)m_{jk}(X_i,\hat{\beta}_{jk}) \right\} , \quad (2.10)$$

where $m_j(X_i,\hat{\beta}_j)$ and $m_{jk}(X_i,\hat{\beta}_{jk})$ are the estimated mean outcomes under treatment pathways A_j and A_jB_k respectively.

This estimator incurs large bias when the first stage treatment assignment, $Z_{ji}^{(A)}$ is not independent of response to that treatment, R_i.

3 Proposed Estimators

3.1 The Hybrid Inverse Probability of Treatment Weighted Estimator

From the causal inference point of view, we are interested in the outcome if the study population had received regime $d(A_j,B_k)$. As discussed earlier, this quantity is $Y_i[d(A_j,B_k)]$. Since all individuals in the sample did not follow this particular regime, $Y_i[d(A_j,B_k)]$ is unobserved for individuals who are treated inconsistent with this regime, and hence missing data techniques can be applied to estimate the parameter of interest, $\mu = \mu[d(A_j,B_k)] = E[Y_i[d(A_j,B_k)]]$. We first define D_{jki}, an indicator that takes the value 1 when the i^{th} individual in the sample is consistent with the treatment regime $d(A_j,B_k)$, and 0 otherwise. More specifically,

$$D_{jki} = Z_{ji}^{(A)}[R_i + (1-R_i) * Z_{ki}^{(B)}] \,. \quad (3.1)$$

The probability of being on the regime $d(A_j,B_k)$, $P(D_{jki}=1)$ is similarly expressed as a composite of the probabilities for treatment assignment and intermediate response:

$$\begin{aligned} P(D_{jki}=1) &= P(Z_{ji}^{(A)}[R_i + (1-R_i) * Z_{ki}^{(B)}] = 1) \\ &= P(Z_{ji}^{(A)}=1)[P(R_i=1|Z_{ji}^{(A)}=1) \\ &\quad + (1 - P(R_i=1|Z_{ji}^{(A)}=1)) * P(Z_{ki}^{(B)}=1|R_i=1,Z_{ji}^{(A)}=1)] \,. \end{aligned}$$

If we were to utilize the probability of being on the regime of interest as a unique weight for each patient we could then

$$\pi^{(D)}_{jk}(X_i,\hat{\gamma}_{jk}) = \pi^{(A)}_{j}(X_i,\hat{\gamma}_{A_j})[\pi^{(R)}_{j|Z^{(A)}_{ji}=1}(X_i,\hat{\gamma}_R) + (1-\pi^{(R)}_{j|Z^{(A)}_{ji}=1}(X_i,\hat{\gamma}_R)) * \pi^{(B)}_{k|Z^{(A)}_{ji}=1,R=1}(X_i,\hat{\gamma}_{B_{jk}})] \,. \quad (3.2)$$

The terms $\pi^{(A)}_{j}(X_i,\hat{\gamma}_{A_j})$, $\pi^{(B)}_{k|Z^{(A)}_{ji}=1,R=1}(X_i,\hat{\gamma}_{B_{jk}})$, and $\pi^{(R)}_{j|Z^{(A)}_{ji}=1}(X_i,\hat{\gamma}_R)$ can be estimated using logistic regression or similar models. An inverse probability of treatment weighted estimator can then be defined as

$$\hat{\mu}_{HIPTW}[d(A_j,B_k)] = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{D_{jki}}{\hat{\pi}^{(D)}_{jk}(X_i,\hat{\gamma_{jk}})}\right\}Y_i \,. \quad (3.3)$$

We will call this new estimator the hybrid inverse probability of treatment weighted (HIPTW) estimator. We call it a hybrid estimator because now we are modeling both treatment and intermediate response, which is a hybrid of G-Computation and IPTW

3.2 The Hybrid Double Robust Estimator

The extension of the HIPTW estimator, (3.3), to a double robust variant is straightforward and analogous to the extension in the single stage setting. The IPTW term is augmented with a model based estimation term that only comes in to play when the probability of being on the regime of interest is incorrectly modeled. We call this the Hybrid Double Robust Estimator and define it as

$$\hat{\mu}_{HDR}[d(A_j,B_k)] = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{D_{jki}}{\hat{\pi}^{(D)}_{jk}(X_i,\hat{\gamma_{jk}})}\,Y_i - \frac{D_{jki}-\hat{\pi}^{(D)}_{jk}(X_i,\hat{\gamma_{jk}})}{\hat{\pi}^{(D)}_{jk}(X_i,\hat{\gamma_{jk}})}\hat{\mu}^{D}_{jk}(X_i,\hat{\beta})\right\}, \quad (3.4)$$

where $\mu^{D}_{jk}(X_i,\beta)$ is the mean of the counterfactual outcome corresponding to being treated with the regime $d(A_j,B_k)$. This new, hybrid two stage double robust estimator has the advantage of being more efficient than the existing two-stage double robust estimator. Additionally it does not require the previously discussed unreasonable assumption of first stage treatment assignment being independent of the outcome of that treatment. Moreover, (3.4) can be viewed as the solution to the estimating equations

$$\psi_i(\theta) = \begin{pmatrix} \psi_{jiA}(X_i, Z^{(A)}_{ji};\gamma_{A_j}) \\ \psi_{iR}(X_i, Z^{(A)}_{ji}, R_i;\gamma_R) \\ \psi_{jkiB}(X_i, Z^{(A)}_{ji}, Z^{(B)}_{ki};\gamma_{B_{jk}}) \\ \psi_{iD}(X_i, D_{jki}, Y_i;\beta) \\ \psi_{iHDR}(X_i, D_{jki}, Y_i;\mu_{HDR}) \end{pmatrix} \quad (3.5)$$

where the ψ functions inside the parenthesis are individual estimating equations for the components of the combined parameter vector,

$$\theta = (\gamma_{A_j}, \gamma_R, \gamma_{B_{jk}}, \beta, \mu_{HDR})' \,, \tag{3.6}$$

used in the estimation of $\mu_{HDR}[d(A_j, B_k)]$. For simplicity we write $\mu[d(A_j, B_k)]$ as μ The first three correspond to logistic regression models for probabilities of first stage treatment assignment and intermediate response while the fourth is the estimating equation for a linear regression model for complete cases. More explicitly,

$$\begin{aligned}
\psi_{jiA}(X_i, Z_{ji}^{(A)}; \gamma_{A_j}) &= X_i \left(Z_{ji}^{(A)} - \frac{1}{1 + e^{-X_i^T \gamma_{A_j}}} \right) \\
\psi_{iR}(X_i, Z_{ji}^{(A)}, R_i; \gamma_R) &= X_i \left(R_i - \frac{1}{1 + e^{-(X_i^T, Z_{ji}^{(A)})\gamma_R}} \right) \\
\psi_{jkiB}(X_i, Z_{ji}^{(A)}, Z_{ki}^{(B)}; \gamma_{B_{jk}}) &= (1 - R_i) X_i \left(Z_{jki}^{(B)} - \frac{1}{1 + e^{-(X_i^T, Z_{ji}^{(A)})\gamma_{B_{jk}}}} \right) \\
\psi_{iD}(X_i, D_{jki}, Y_i; \beta) &= R_{jki}^{*} (Y_i - \mu_{jk}^{D}(X_i, \hat{\beta})) \,.
\end{aligned}$$

The final equation is the estimating equation corresponding to the hybrid double robust estimator itself, namely,

$$\psi_{i\mu}(X_i, D_{jki}, Y_i; \mu_{HDR}) = \frac{D_{jki}\, Y_i}{\pi_{jk}^{(D)}(X_i, \gamma_{jk})} - \frac{D_{jki} - \pi_{jk}^{(D)}(X_i, \gamma_{jk})}{\pi_{jk}^{(D)}(X_i, \gamma_{jk})} \mu_{jk}^{D}(X_i, \beta) - \mu[d(A_j, B_k)] \,.$$

Defining the hybrid double robust estimator as a solution to the above set of estimating equations allows us to classify it as an M-estimator. It follows then that it is asymptotically normal, as well as consistent (Stefanski and Boos, 2002). An estimator of the variance of the Hybrid Double Robust estimator can be found with the sandwich estimator of the variance $V_n = A_n(\hat{\theta})^{-1} B_n(\hat{\theta}) A_n(\hat{\theta})^{-1^T}/n$, where

$$A_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} -\frac{\delta \psi_i(\theta)}{\delta \theta} = \frac{1}{n} \sum_{i=1}^{n} \begin{pmatrix}
-\frac{\delta \psi_{jiA}}{\delta \gamma_{A_j}} & 0 & 0 & 0 & 0 \\
0 & -\frac{\delta \psi_{iR}}{\delta \gamma_R} & 0 & 0 & 0 \\
0 & 0 & -\frac{\delta \psi_{jkiB}}{\delta \gamma_{B_j k}} & 0 & 0 \\
0 & 0 & 0 & -1 & 0 \\
-\frac{\delta \psi_{i\mu}}{\delta \gamma_{A_j}} & -\frac{\delta \psi_{i\mu}}{\delta \gamma_R} & -\frac{\delta \psi_{i\mu}}{\delta \gamma_{B_{jk}}} & -\frac{\delta \psi_{i\mu}}{\delta \beta} & -1
\end{pmatrix}$$

and

$$B_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} \psi_i(\theta) \psi_i(\theta)^T,$$

where the derivatives in the matrix $A_n(\theta)$ are defined as

$$\frac{\delta\psi_{jiA}}{\delta\gamma_{A_j}} = X_i e^{-X_i^T\gamma_{A_j}}\left(\frac{1}{1+e^{-X_i^T\gamma_{A_j}}}\right)^2 X_i^T$$

$$\frac{\delta\psi_{jkiB}}{\delta\gamma_{B_{jk}}} = X_i e^{-X_i^T\gamma_{B_{jk}}}\left(\frac{1}{1+e^{-X_i^T\gamma_{B_k}}}\right)^2 X_i^T$$

$$\frac{\delta\psi_{iR}}{\delta\gamma_R} = X_i e^{-X_i^T\gamma_R}\left(\frac{1}{1+e^{-X_i^T\gamma_R}}\right)^2 X_i^T \qquad \frac{\delta\psi_{i\mu}}{\delta\gamma_{A_j}} = \frac{\pi^D_{\gamma_{A_j}} R_i * (Y_i - \mu^D_{jk})}{(\pi^D)^2}$$

$$\frac{\delta\psi_{i\mu}}{\delta\gamma_{B_{jk}}} = \frac{\pi^D_{\gamma_{B_{jk}}} R_i * (Y_i - \mu^D_{jk})}{(\pi^D)^2} \qquad \pi^D_{\gamma_{B_{jk}}} = \frac{\delta\pi^{(D)}_{jk}(X_i,\gamma_{jk})}{\delta\gamma_{B_jk}}$$

$$\frac{\delta\psi_{i\mu}}{\delta\gamma_R} = \frac{\pi^D_{\gamma_R} R_i * (Y_i - \mu^D_{jk})}{(\pi^D)^2} \qquad \pi^D_{\gamma_{A_j}} = \frac{\delta\pi^{(D)}_{jk}(X_i,\gamma_{jk})}{\delta\gamma_{A_j}}$$

$$\frac{\delta\psi_{i\mu}}{\delta\beta} = \frac{R_i - \pi_D}{\pi_D} \qquad \pi^D_{\gamma_R} = \frac{\delta\pi^{(D)}_{jk}(X_i,\gamma_{jk})}{\delta\gamma_R}.$$

A similar approach can be used to find an estimator of the variance for the hybrid inverse probability of treatment weighted estimator, the difference being that the $A_n(\theta)$ and $B_n(\theta)$ matrices have smaller dimensions.

4 Simulations

We based our simulation on the one found in Kang and Schafer (2007) with a few notable exceptions. For each subject we generate 8 independent identically distributed standard normal random variables, which we will call X_a, X_b, X_c, X_d, X_e, X_f, X_g, and X_h. We assume a two stage design, with non-responders from the first stage being rerandomized to the second stage. Without loss of generality we will be dropping the j and k subscript, concerning ourselves only with the regime where $Z^{(A)}_{ji} = 1$ and $Z^{(B)}_{ki} = 1$. The treatment probabilities for each stage are

$$\pi^{(A)}_j = expit(\gamma_{10} + \gamma_{11} * X_e + \gamma_{12} * X_f + \gamma_{13} * X_g + \gamma_{14} * X_h) \tag{4.1}$$

and

$$\pi^{(B)}_{jk} = expit(\gamma_{20} + \gamma_{21} * X_e + \gamma_{22} * X_f + \gamma_{23} * X_g + \gamma_{24} * X_h + \gamma_{25} * Z^{(A)}_{ji}) \,, \tag{4.2}$$

where $expit$ is the inverse logit function.

With probability of response to the first stage being defined as

$$R \sim BERNOULLI(\tau_0 + \tau_1 * Z^{(A)}_{ji}) \,. \tag{4.3}$$

The outcome variable, Y, is then written as

$$Y = \beta_0 + \beta_1 * Z_A + \beta_2 * Z_B + \beta_3 * X_a + \beta_4 * X_b + \beta_5 * X_c + \beta_6 * X_d + \beta_7 * R + \beta_8 * Z^{(A)}_{ji} * Z^{(B)}_{ki} + e \,, \tag{4.4}$$

where e $\sim N(0,1)$ are normally distributed error terms. Now the observed covariates are

$$\begin{aligned} X_1 &= e^{\frac{X_a}{2}}, \quad X_2 = \frac{X_b}{1+e^{X_a}} + 10, \quad X_3 = \left(\frac{X_a \times X_c}{25} + .6\right)^3 \\ X_4 &= (X_b + X_d + 20)^2, \quad X_5 = e^{\frac{X_e}{2}}, \quad X_6 = \frac{X_f}{1+e^{X_e}} + 10 \\ X_7 &= \left(\frac{X_e \times X_g}{25} + .6\right)^3, \quad X_8 = (X_f + X_h + 20)^2 \, . \end{aligned} \tag{4.5}$$

Since it would be unreasonably difficult for any investigator to correctly identify the correct relationship between observed covariates and either outcome or treatment we assume the following incorrect models could be specified:

$$Y = \xi_{10} + \xi_{11} * X_1 + \xi_{12} * X_2 + \xi_{13} * X_3 + \xi_{14} * X_4 + \xi_{15} * Z_{ji}^{(A)} + e \, , \tag{4.6}$$

which is the outcome model for those who did respond to the initial treatment.

$$\begin{aligned} Y = \xi_{20} + \xi_{21} * X_1 + \xi_{22} * X_2 + \xi_{23} * X_3 \\ + \xi_{24} * X_4 + \xi_{25} * Z_{ji}^{(A)} + \xi_{26} * Z_{ki}^{(B)} + \xi_{27} * Z_{ji}^{(A)} * Z_{ki}^{(B)} + e \, , \end{aligned} \tag{4.7}$$

which is the outcome model for those who did not respond to the initial treatment and were rerandomized to a second stage treatment.

$$\pi_j^{(A)} = \eta_{11} * X_5 + \eta_{12} * X_6 + \eta_{13} * X_7 + \hat{\eta_{14}} * X_8 \tag{4.8}$$

and

$$\pi_{jk}^{(B)} = \eta_{21} * X_5 + \eta_{22} * X_6 + \eta_{23} * X_7 + \hat{\eta_{24}} * X_8 + \hat{\eta_{25}} * Z_{ji}^{(A)} \, , \tag{4.9}$$

which are logistic regressions of the treatment assignments on the covariates with ξ and η as their estimated coefficients.

We perform a series of 10,000 Monte Carlo simulations in order to asses the validity of our estimator under circumstances where one or both of the types of models are miss-specified. For the first simulation the true value of the estimand, the mean counterfactual outcome

$$\begin{aligned} \mu\{d(A_1, B_1)\} &= E\{Y[d(A_1, B_1)]\} = 142.4, \\ \beta &= (110, 24.7, 13.7, 124.7, 73.7, 23.7, 50.7, 20.2, -10.7)' \\ \tau &= (.2, -.7)' \\ \gamma &= (\gamma_{10}, \gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14}, \gamma_{20}, \gamma_{21}, \gamma_{22}, \gamma_{23}, \gamma_{24}, \gamma_{25})' \\ &= (-.3, -.75, -.2, .5, 1, .2, -.2, -.3, .4, -.5, -.7)'. \end{aligned}$$

In the second simulation, created to more closely mimic the STAR*D data, the true value of the

estimand is

$$\begin{aligned}\mu\{Y[d(A,B)]\} &= E\{d(A,B)\} = 7.4\\ \beta &= (-0.932, -0.137, 1.51, 0.045, 0.313, 0, 0, 5.09, -0.951)'\\ \tau &= (-0.559, 0.711)'\\ \gamma &= (0, -0.457, -0.135, 0, 0, 0, -0.074, -0.041, 0, 0, .500)'.\end{aligned}$$

We also varied sample size in both simulations, at $n = 200$ and $n = 500$ for the first simulation and $n = 2000$ for the second, as well as examining the estimators under both standard normal and t-distribution error structures in the first simulation and standard normal error structures in the second.

4.1 Simulation Results

Table 1: Simulation results for n=200 and n=500, Standard Normal Errors, $\mu\{d(A_1, B_1)\} = 142.4$, G-Comp = G-Computation, IPTW = Inverse Probability of Treatment Weighted, DR = Double Robust, HIPTW = Hybrid IPTW, HDR = Hybrid DR, MCSE = Monte Carlo Standard Error, MSE = Mean-Squared Error.

Model		n=200			n=500		
Treatment Outcome	Method	Bias (%)	MCSE	Relative Efficiency	Bias	MCSE	Relative Efficiency
Correct Correct	G-Comp	0.4	21.4	3.2	0.7	13.3	2.8
	IPTW	-0.5	38.1	9.9	0.6	24.6	9.2
	DR	4.2	20.2	3.1	3.0	12.5	2.7
	HIPTW	-2.9	31.5	3.9	-2.4	19.2	5.9
	HDR	-2.6	11.2	1	-2.4	7.1	1
Incorrect Correct	G-Comp	0.4	21.4	3.2	0.7	13.3	2.29
	IPTW	4.4	63.6	27.4	8.8	144.0	252.7
	DR	0.9	20.1	2.8	-1.9	15.7	3.2
	HIPTW	2.1	66.6	29.9	6.1	156.0	294.6
	HDR	-2.5	11.3	1	-2.3	8.1	1
Correct Incorrect	G-Comp	0.4	21.4	2.3	0.7	13.30	2.2
	IPTW	0.5	38.1	7.2	0.6	24.6	7.1
	DR	4.2	20.2	2.2	3.0	12.5	2.1
	HIPTW	-2.9	31.5	5.0	-2.4	19.2	4.5
	HDR	-1.6	13.7	1	-1.9	8.5	1
Incorrect Incorrect	G-Comp	0.4	21.4	478	0.7	13.3	0.2
	IPTW	4.4	63.6	13.1	8.8	144.0	17.9
	DR	3.3	36.9	4.5	-0.5	17.4	0.3
	HIPTW	2.1	66.6	14.3	6.1	156.0	20.9
	HDR	-1.6	17.1	1	-1.5	33.9	1

Table 2: Simulation results for n=200 and n=500, t-Distribution with df=5 Errors, $\mu\{d(A_1, B_1)\} = 142.4$, G-Comp = G-Computation, IPTW = Inverse Probability of Treatment Weighted, DR = Double Robust, HIPTW = Hybrid IPTW, HDR = Hybrid DR, MCSE = Monte Carlo Standard Error, MSE = Mean-Squared Error.

Model		n=200			n=500		
Treatment Outcome	Method	Bias (%)	MCSE	MSE	Bias (%)	MCSE	MSE
Correct Correct	G-Comp	0.5	21.4	3.3	0.7	13.4	2.9
	IPTW	-0.1	46.4	15.0	0.311	23.5	8.5
	DR	4.6	37.6	10.2	3.0	12.4	2.7
	HIPTW	-3.0	31.5	7.1	-2.6	18.8	5.7
	HDR	-2.4	11.1	1	-2.4	7.0	1
Incorrect Correct	G-Comp	0.5	21.4	2.9	0.7	13.4	0.03
	IPTW	5.3	136.0	114.1	-9.5	2390.0	747.7
	DR	0.6	37.2	8.7	-1.7	15.3	0.03
	HIPTW	2.5	75.5	35.4	-17.1	3050.0	121.9
	HDR	-2.3	11.8	1	-2.9	86.7	1
Correct Incorrect	G-Comp	0.5	21.4	2.3	0.7	13.4	2.2
	IPTW	-0.1	46.4	10.7	0.3	23.5	6.4
	DR	4.6	37.6	7.3	3.0	12.4	2.1
	HIPTW	-3.0	31.5	5.1	-2.6	18.8	4.3
	HDR	-1.5	13.7	1	-2.0	8.5	1
Incorrect Incorrect	G-Comp	0.5	21.4	0.8	0.7	13.4	$<$0.01
	IPTW	5.3	136.0	30.6	-9.5	2390.0	14.8
	DR	3.7	50.9	4.4	-0.3	16.8	$<$0.01
	HIPTW	2.5	75.5	9.5	-17.1	3050.0	24.1
	HDR	-1.4	24.1	1	-6.1	620.0	1

Results from Scenario 1 are presented in Table 1. When both treatment and outcome models are correct, all the estimators except for the existing DR estimator are approximately unbiased with relative biases below 3%. For the existing DR estimator, this bias is 4.2% versus a bias of -2.6% for the HDR estimator. The G-Comp and IPTW estimators, as expected, have the smaller biases compared to others (since both models are correct), 0.4% and -0.5% respectively, while the HIPTW estimator has a larger bias (-2.9%). Likewise, all of the estimators except for the IPTW and HIPTW have Monte Carlo standard errors under 30. For this scenario, the IPTW estimator has a MCSE of 38.1 compared to the HIPTW's MCSE of 31.5. The HDR estimator performs the best with a MCSE of 11.2 versus 20.2 for the existing DR estimator and 21.4 for the G-Computation estimator. As in the case of the MCSE, the MSE of the existing IPTW and HIPTW estimators is much larger than the other estimators, at 9.9 and 3.9 times higher than the HDR estimator's MSE respectively. The HDR estimator, with an MSE of 150, also outperforms the existing DR and G-Comp estimators, who are 3.1 and 3.2 times higher than that respectively. When only the outcome models are incorrectly specified the estimators maintain their relationships to one another in terms of bias, Monte Carlo

Table 3: Performance of the sandwich variance estimator and corresponding confidence intervals for the HDR estimator with X_8, n=200 and n= 500, Standard Normal Errors, $\mu[d(A_1, B_1)] = 142.4$

Model	n=200			n=500		
Treatment Outcome	Sandwich Est. SE	MCSE	Coverage (%)	Sandwich Est. SE	MCSE	Coverage (%)
Correct Correct	11.1	11.2	93.7	7.0	7.1	92.0
Incorrect Correct	128.1	11.3	99.8	63.5	8.1	99.8
Correct Incorrect	13.5	13.7	95.8	8.5	8.5	93.1
Incorrect Incorrect	610.2	17.1	100	310.0	33.9	100

standard error, and mean squared error.

However, we see some of these relationships change when only the treatment models are incorrectly specified. The DR estimator in this case has a lower bias than the HDR estimator at 0.9% and -2.5% respectively. The HIPTW estimator also has lower bias, at 2.1%, than both the HDR estimator and the existing IPTW estimator, which has a bias of 4.4%. Also, in this case, the relationship between the existing IPTW and HIPTW estimators is reversed in terms of both MCSE and MSE. The IPTW estimator has a lower MCSE and MSE, 5.6 and 27.4 times higher than the HDR estimator respectively. While the HIPTW estimator has an MCSE of 66.6 and an MSE 29.9 times higher than that of the HDR estimator. The other estimators keep their relationships to one another and maintain Monte Carlo standard errors under 30 and mean squared errors around 3 times higher than that of the HDR estimator.

When both treatment and outcome models are incorrectly specified the relationships between these estimators change slightly from when only the treatment models were incorrectly specified. In terms of relative bias the IPTW and DR estimators are now both over 3% with 4.4% and 3.1% respectively. In this case, the HIPTW estimator is less biased, at 2.1%, than both the IPTW and DR estimators. Additionally the DR estimator has an MCSE over 30, at 36.9, and an MSE 4.5 times that of the HDR. G-Comp estimator has a lower MSE, at only 1.5 times that of the HDR estimator. In this case, only the HDR estimator has a lower MSE than the G-Comp estimator.

In summary, Table 1 shows us some interesting dynamics between the estimators. When the treatment models are correctly specified, the IPTW estimator is less biased, has a higher Monte Carlo standard error, and a lower MSE than the new Hybrid IPTW estimator. This result, however, is reversed when the treatment model is miss-specified. The Hybrid Double Robust estimator has the lowest Monte Carlo standard errors and MSE of any estimator regardless of model miss-specification at this sample size of $n = 200$. The DR estimator has lower bias than the HDR estimator only when the treatment models are incorrectly specified but the outcome models are not. G-comp has the lowest bias of any of the methods but it is important to note that only the outcome models for

Table 4: Performance of the sandwich variance estimator and corresponding confidence intervals for the HDR estimator with X_8, n=200 and n= 500, t-Distribution with df=5 Errors, $\mu[d(A_1, B_1)] = 142.4$

Model	n=200			n=500		
Treatment Outcome	Sandwich Est. SE	MCSE	Coverage (%)	Sandwich Est. SE	MCSE	Coverage (%)
Correct Correct	11.1	11.2	93.7	7.0	7.1	92.5
Incorrect Correct	12.0	11.2	95.5	7.4	8.1	93.4
Correct Incorrect	13.5	13.7	95.8	8.5	8.5	93.1
Incorrect Incorrect	20.5	14.2	98.0	10.8	9.19	96.9

pathway endpoints are incorrectly specified, not the models for intermediate response. G-comp and IPTW produce estimators with biases of similar magnitudes when both treatment and outcome models are correctly specified. In Table 1, where $n = 500$ and when both the outcome and treatment models are incorrectly specified the HDR estimator performs worse than the DR estimator in terms of bias (-1.5% versus -.05%), MCSE (33.9 versus 17.4), and with an MSE 0.3 times that of the HDR estimator.

Table 2 are the results for running the simulation at $n = 200$ and $n = 500$ with error terms being generated from a student's t-distribution with 5 degrees of freedom. It is similar to Table 1 with the exception that in this case when both outcome and treatment models are incorrectly specified the HDR estimator is inferior to the G-comp estimator in terms of MCSE (24.1 versus 21.4) with an MSE that is 1.25 times higher. One of the Monte Carlo simulation runs for $n = 500$ contained a single observation that, under the incorrectly specified model for estimating probability of receiving treatment A, received a very small probability of receiving its actual treatment assignment. This in turn caused that observation to receive an abnormally high weight. The estimators utilizing inverse probability weighting for this run suffered from increased estimated biases, Monte Carlo standard errors, and MSEs as a result of this single run.

Table 3 shows a comparison between the standard error computed using the sandwich estimator of the variance and the Monte Carlo standard error at $n = 200$. The table shows that when the treatment models are correctly specified the sandwich estimator produces close but slightly under estimated estimators of the standard error. When both models are correctly specified the average of the sandwich variance estimator is 11.1 versus the MCSE of 11.2, and when only the outcome is incorrectly specified the these values are respectively 13.5 and 13.7. When the treatment models are incorrectly specified then the sandwich estimator of the variance appears not to be very accurate at all. With only the treatment models incorrectly specified the sandwich estimator is 128.1 versus the MCSE of 11.3. When both models are incorrectly specified the sandwich estimator becomes even

Table 5: Simulation results for n=2000, Standard Normal Errors, $\mu\{d(A_1, B_1)\} = 7.4$, G-Comp = G-Computation, IPTW = Inverse Probability of Treatment Weighted, DR = Double Robust, HIPTW = Hybrid IPTW, HDR = Hybrid DR, MCSE = Monte Carlo Standard Error, MSE = Mean-Squared Error.

Treatment Model Outcome Model	Method	Bias	MCSE	Relative Efficiency
Correct Correct	G-Comp	-0.002	0.135	0.291
	IPTW	0.007	0.146	0.317
	DR	9.78	0.116	1.24
	HIPTW	-8.00	0.158	1.01
	HDR	-8.04	0.150	1
Incorrect Correct	G-Comp	-0.002	0.135	0.283
	IPTW	-0.06	0.158	0.339
	DR	8.10	0.157	1.001
	HIPTW	-6.56	0.191	0.857
	HDR	-8.04	0.161	1
Correct Incorrect	G-Comp	-0.002	0.135	0.290
	IPTW	0.007	0.146	0.317
	DR	9.78	0.116	1.24
	HIPTW	-8.00	0.158	1.01
	HDR	-8.03	0.151	1
Incorrect Incorrect	G-Comp	-0.00213	0.135	0.283
	IPTW	-0.006	0.158	0.338
	DR	8.10	0.157	0.998
	HIPTW	-6.56	0.191	0.856
	HDR	-8.04	0.161	1

less accurate with 610.2 versus the MCSE of 17.1. Table 3 corroborates this by showing similar discrepancies under treatment model miss-specification at $n = 500$. The coverage at both sample sizes are suboptimal when the treatment model is correctly specified, and highly conservative when incorrect. A close examination of the simulations revealed that the elements of the matrix $A_n(\hat{\theta})$ corresponding to X_8, a variable in the miss-specified treatment models, were abnormally large and were contributing the majority of the estimated error.

Table 4 shows the results of the simulation with X_8 removed from the miss-specified treatment models at $n = 200$ and $n = 500$. When the treatment models are incorrectly specified and the outcome models are not, the sandwich estimator is slightly over-estimating the variance at $n = 200$, with a value of 12.0 versus the MCSE of 11.2, and slightly under-estimating at $n = 500$, with a value of 7.4 versus the MCSE of 8.1. When both the treatment models and the outcome models are incorrectly specified the standard error derived from the sandwich estimator of the variance slightly over-estimates the standard error. The over-estimation is more severe at the smaller sample size with the average sandwich estimator value of 20.5 versus the MCSE of 14.2. At $n = 500$ this is less concerning with corresponding values 10.8 and 9.19 respectively. At $n = 200$ the coverage probabilities when at least one model is correctly specified are close to 95% while at $n = 500$ they

Table 6: Performance of sandwich variance estimator n=2000, Standard Normal Errors, $\mu\{d(A_1, B_1)\} = 7.4$, HDR = Hybrid Double Robust Estimator

Treatment Model Outcome Model	Method	Sandwich Est. SE	MCSE	Coverage (%)
Correct Correct	HDR	0.146	0.150	85.0
Incorrect Correct	HDR	0.138	0.161	80.1
Correct Incorrect	HDR	0.146	0.151	83.8
Incorrect Incorrect	HDR	0.156	0.161	85.8

become slightly suboptimal. When both models are miss-specified the coverage at both sample sizes are larger then 95%.

Finally, in Tables 5 and 6 we see that in the second simulation the HDR estimator performs worse than the other estimators. There is an increase in bias, likely caused by the smaller ratio of intercept to treatment effect coefficients in the outcome model. The stronger dependence of first stage treatment on outcome is problematic with the Hybrid estimators in such a scenario. The lower coverage for HDR in the the second simulation is due to the higher bias of the estimator.

4.2 Summary of the Simulation Results

In this simulation study we have examined several current methods for estimating the mean outcome under a dynamic treatment regime along with the two newly introduced hybrid estimators. First, the new Hybrid estimators are more efficient than their non-hybrid counterparts. Also at low sample sizes, with the exception of when only the outcome is correctly specified, the new hybrid double robust estimator is both less biased and has a smaller mean squared error then the classic double robust estimator. The hybrid inverse probability of treatment weighted estimator is more biased than its non-hybrid counterpart under correct model specification but considerably less biased when the treatment model is incorrectly specified. The hybrid inverse probability weighted estimator has a lower mean squared than its non-hybrid counterpart regardless of correct treatment model specification.

At larger sample sizes, the hybrid double robust estimator is less biased, more efficient, and has a lower mean squared error than the non-hybrid double robust estimator but only when the treatment model is correctly specified. Under correct treatment model specification, the hybrid inverse probability of treatment weighted estimator is more efficient and has a lower mean squared standard error than its counterpart but does not maintain that advantage under treatment model miss-specification at higher sample sizes.

With non-standard error terms generated using the student's t-distribution with 5 degrees of

freedom, all of the advantages of the hybrid estimators at low sample sizes are still present but at larger sample sizes the hybrid methods faults are magnified in the case of treatment model miss-specification. This is the result of the t-distribution being very tail heavy in terms of its distribution producing extreme weights when the estimated treatment probability is poorly constructed and lead to unstable estimates. One solution to this would be to truncate very high or very low weights, which is a common practice in inverse probability of treatment weighting.

The sandwich estimator of the variance appears to be unbiased but unstable when the treatment model is miss-specified.

5 The STAR*D Trial

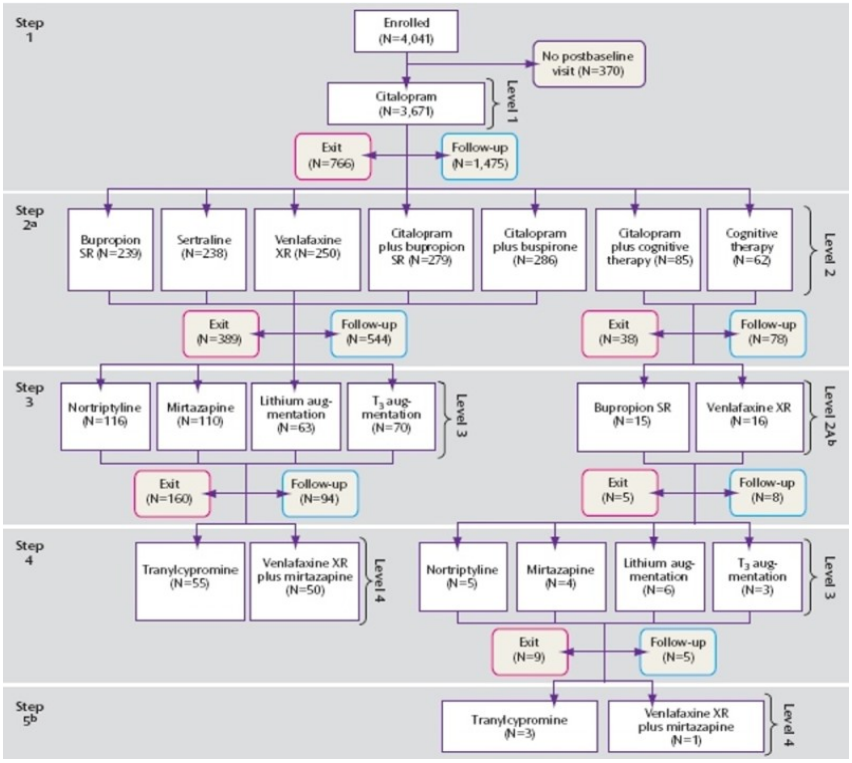

Figure 1: STAR*D Stages of Treatment (Warden et al., 2007)

The Sequenced Treatment Alternatives to Relieve Depression, or STAR*D, Trial was a 7 year multistage trial designed to test different regimes of depression treatments. The trial contained 4,041 patients between the ages of 18 and 75 who had received a clinical diagnosis of non-psychotic major depressive disorder under the DSM-IV checklist. A patient that is considered to respond to treatment will continue taking that treatment, while those that do not respond are re-randomized

Figure 2: Forest Plot of STAR*D Dynamic Treatment Regimes; QIDS-16 Score of 5 = Remission; BUP - bupropion sustained release; BUS - buspirone; CIT - citalopram; Li - lithium; MIRT - mirtazapine; NTP - nortriptyline; SER - sertraline; THY - triiodothyronine

Table 7: STAR*D Dynamic Treatment Regimes

Regime	n	Mean (QIDS-16)	95% Confidence Interval
d(BUP,MIRT)	95	7.4	(7.3, 7.5)
d(BUP,NTP)	93	7.8	(7.7, 7.9)
d(BUP,BUP+LI)	85	7.7	(7.6, 7.9)
d(BUP,BUP+THY)	73	6.2	(5.9, 6.5)
d(SER,MIRT)	117	7.4	(7.4, 7.5)
d(SER,NTP)	108	6.7	(6.6, 6.8)
d(SER,SER+LI)	97	6.3	(6.2, 6.4)
d(SER,SER+THY)	96	6.1	(6.0, 6.2)
d(VEN,MIRT)	120	7.4	(7.3, 7.5)
d(VEN,NTP)	119	6.7	(6.6, 6.7)
d(VEN,VEN+LI)	108	6.9	(6.8, 7.0)
d(VEN,VEN+THY)	113	6.2	(6.2, 6.3)
d(CIT+BUP,MIRT)	157	6.4	(6.3, 6.5)
d(CIT+BUP,NTP)	163	6.7	(6.7, 6.8)
d(CIT+BUP,CIT+LI)	161	6.5	(6.4, 6.6)
d(CIT+BUP,CIT+THY)	167	5.9	(5.9, 6.0)
d(CIT+BUS,MIRT)	153	7.7	(7.7, 7.8)
d(CIT+BUS,NTP)	160	7.2	(7.1, 7.3)
d(CIT+BUS,CIT+LI)	145	8.1	(8.0, 8.2)
d(CIT+BUS,CIT+THY)	151	6.7	(6.7, 6.8)

to a treatment within a treatment-grouping of their own choice. Although the trial was designed with four and a half stages, we will only be looking at the 2nd and 3rd stage treatments. This is because everyone received the same treatment in the first stage, an SSRI called citalopram, and if we included the 4th stage into our analysis our regime sample sizes would be too small to make any reasonable inferences. We will also be excluding patients that received cognitive therapy in the second stage. They have the possibility of being assigned to an extra half stage, which would make analysis difficult in ways similar to if we had included the 4th stage in our analysis. So our analysis will center on comparing regimes for follow up medications for patients with depression that are non-responsive to SSRIs. For that reason stage 2 and 3 treatments will be henceforth be represented by 1st and 2nd stage notations under our conventions. So a response to Stage 2 would be indicated by R_1, the outcome model for Stage 3 would be denoted m_{jk}, and so on. A visual representation of the trial can be seen in Figure 1. Thus there would be a total of 20 dynamic treatment regimes to compare, namely, d(a, b), where $b \in$ (MIRT, NTP, a+LI, a+THY), with $a \in$ (BUP, SER, VEN, CIT+BUP, CIT+BUS).

Response is defined as either a 50% reduction in QIDS-16 (Quick Inventory of Depressive Symptomatology) self reported score, which the investigators call response, or a QIDS-16 score lower than 5, which the investigators call remission, at the end of the stage. The QIDS-16 score is also the outcome measure we will be looking at in our analysis, rather than the HRSD-17 (Hamilton Rating Scale of Depression) score which the original investigators used. This is because QIDS-16

was recorded at every visit while HRSD-17 was recorded upon exiting or finishing the trial, QIDS-16 scores would be available for more patients than the HRSD-17 was for. It should be noted that the QIDS-16 remission rates were generally higher than HRSD-17 rates, mainly because the study counted not having an exit HRSD-17 score as non-response. We will also differ from the study in that we will keep our overall outcome continuous rather than transform it to being binary. Transforming a continuous measure into a binary one is always accompanied by a loss of information(Altman and Royston, 2006). As with the original analysis by the investigators we assume, one could argue wrongly, that dropout from the study represents non-response and we use last observation carried forward as the final outcome.

The treatment probabilities for each stage were estimated using logistic regression modeling probability of receiving treatment of interest on the covariates Burden of Side Effects Rating (Wisniewski et al., 2006) and change in QIDS-16 score that were measured during the previous stage. The outcomes were linearly modeled using our previously described β estimating technique with the covariates being age and QIDS-16 score at the beginning of Stage 2.

Applying our new, more efficient double robust estimator to the data yields the estimated mean QIDS-16 scores under various treatment regimes depicted in Table 7 and Figure 2.

By this analysis, the best follow up treatment if citalopram is not effective is found to be either sertraline or citalopram combined with bupropion sustained release. Additionally the best regimes involve the third stage treatment of supplementing the second stage treatment with triiodothyronine. The best overall regime appears to be 'If non-response to citalopram augment it with bupropion sustained release and if non-response to that switch to supplementing citalopram with triiodothyronine.'

6 Discussion

In this paper, we have introduced a new way of constructing estimators for two-stage dynamic treatment regimes, creating both a hybrid inverse probability of treatment weighted estimator and a hybrid double robust estimator. Through simulations we have shown that these estimators generally perform better than their non-hybrid counterparts by the metrics of bias, standard error, and mean standard error. Finally we have applied the new hybrid double robust estimator to the STAR*D data set in order to estimate the treatment effects of different two stage dynamic treatment regimes for patients with non-psychotic major depressive disorder who do not respond to SSRIs.

One limitation of the methodology described in this paper is that it discards data for whom the outcome data or treatment data are missing. This may result in bias, or efficiency loss as patients in trials can drop out for various reasons. Future work could potentially include an estimator that, in addition to dealing with the missing data problem of not seeing the outcome under the regime of interest for all patients, addresses missing data due to dropout.

Acknowledgment

Andrew Topp's research, while at the University of Pittsburgh, was funded by a Patient-Centered Outcomes Research Institute (PCORI) Project Award (ME-1306-03827) [Disclaimer: All statements

in this report, including its findings and conclusions, are solely those of the author and do not necessarily represent the views of the PCORI, its Board of Governors, or Methodology Committee]. The authors acknowledge the comments from the editors and the reviewers which helped improve the manuscript.

References

Altman, D. G. and Royston, P. (2006), "The cost of dichotomising continuous variables," *BMJ*, 332, 1080.

Bai, X., Tsiatis, A. A., and O'Brien, S. M. (2013), "Doubly-Robust Estimators of Treatment-Specific Survival Distributions in Observational Studies with Stratified Sampling," *Biometrics*, 69, 830–839.

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

Bembom, O. and van der Laan, M. J. (2007), "Statistical methods for analyzing sequentially randomized trials," *Journal of the National Cancer Institute*, 99, 1577–1582.

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

Carpenter, J. R., Kenward, M. G., and Vansteelandt, S. (2006), "A comparison of multiple imputation and doubly robust estimation for analyses with missing data," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169, 571–584.

Holland, P. W. (1986), "Statistics and causal inference," *Journal of the American statistical Association*, 81, 945–960.

Horvitz, D. G. and Thompson, D. J. (1952), "A generalization of sampling without replacement from a finite universe," *Journal of the American statistical Association*, 47, 663–685.

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

Lunceford, J. K., Davidian, M., and Tsiatis, A. A. (2002), "Estimation of Survival Distributions of Treatment Policies in Two-Stage Randomization Designs in Clinical Trials," *Biometrics*, 58, 48–57.

Matthay, K. K., Villablanca, J. G., Seeger, R. C., Stram, D. O., Harris, R. E., Ramsay, N. K., Swift, P., Shimada, H., Black, C. T., Brodeur, G. M., et al. (1999), "Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-cis-retinoic acid," *New England Journal of Medicine*, 341, 1165–1173.

Murphy, S. A. (2003), "Optimal dynamic treatment regimes," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65, 331–355.

Robbins, J. M., Hernan, M. A., and Brumback, B. (2000), "Marginal structural models and causal inferences in epidemiology," *Epidemiology*, 11, 550–60.

Robins, J. (1986), "A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect," *Mathematical Modelling*, 7, 1393–1512.

Robins, J. M. (2000), "Robust estimation in sequentially ignorable missing data and causal inference models," in *Proceedings of the American Statistical Association*, vol. 1999, pp. 6–10.

Rotnitzky, A., Lei, Q., Sued, M., and Robins, J. M. (2012), "Improved double-robust estimation in missing data and causal inference models," *Biometrika*, 99, 439–456.

Stefanski, L. A. and Boos, D. D. (2002), "The calculus of M-estimation," *The American Statistician*, 56, 29–38.

Trivedi, M. H., Kocsis, J. H., Thase, M. E., Morris, D. W., Wisniewski, S. R., Leon, A. C., Gelenberg, A. J., Klein, D. N., Niederehe, G., Schatzberg, A. F., et al. (2008), "REVAMP—research evaluating the value of augmenting medication with psychotherapy: rationale and design," *Psychopharmacol Bull*, 41, 5–33.

Tsiatis, A. A., Davidian, M., and Cao, W. (2011), "Improved doubly robust estimation when data are monotonely coarsened, with application to longitudinal studies with dropout," *Biometrics*, 67, 536–545.

Wahed, A. S. and Tsiatis, A. A. (2004), "Optimal Estimator for the Survival Distribution and Related Quantities for Treatment Policies in Two-Stage Randomization Designs in Clinical Trials," *Biometrics*, 60, 124–133.

Warden, D., Rush, A. J., Trivedi, M. H., Fava, M., and Wisniewski, S. R. (2007), "The STAR* D Project results: a comprehensive review of findings," *Current psychiatry reports*, 9, 449–459.

Wisniewski, S. R., Rush, A. J., Balasubramani, G., Trivedi, M. H., Nierenberg, A. A., Investigators, S., et al. (2006), "Self-rated global measure of the frequency, intensity, and burden of side effects," *Journal of Psychiatric Practice®*, 12, 71–79.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2013), "Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions," *Biometrika*, 681–694.

Received: July 17, 2018

Accepted: August 26, 2018