POLICY LEARNING WITH RARE OUTCOMES

JULIA HATAMYAR AND NOEMI KREIF

Abstract. In settings with heterogeneous policy impacts, observational data can be used to learn optimal policy assignment rules under unconfoundedness. One recent approach (Athey and Wager (2021)) searches over shallow decision trees, using estimates of double-robust (DR) scores obtained via machine learning, to maximise the expected value of a particular treatment rule. In many applications, policy impacts are commonly evaluated for rare outcomes, where it may be challenging to obtain estimates that fully capture treatment effect heterogeneity. Little is known about how ML algorithms perform in this setting. This paper therefore contrasts various methods for obtaining DR-scores, using simulated data with varying degrees of overlap quality and treatment effect heterogeneity, and reports their performance in terms of errors in estimated average and heterogeneous treatment effects, and the value of learned policy rules for rare outcomes. The methods are applied to a case study with the goal of reducing infant mortality through improved targeting of subsidized health insurance in Indonesia. Although the methods perform similarly in settings with normal outcome prevalence, there are clear winners when outcomes are rare. Bayesian additive regression trees (BART) perform particularly well in situations where the sample size is relatively low compared to the number of covariates.

1. Introduction

In the health and social sciences, there is increasing interest in using observational data to learn optimal policy assignment rules that map an individual's covariate profile to a treatment decision (Murphy 2003; Manski 2004; Hirano and Porter 2009; Kitagawa and Tetenov 2018). These rules can be optimal in the sense that they maximise expected outcomes, such as health gains. The problem of learning optimal policy assignment rules is linked

Date: May 24, 2022.
Julia Hatamyar, Corresponding Author e-mail: julia.hatamyar@york.ac.uk. All `R` code used in this paper is available at `github.com/jhatamyar/OptPolicyRare`.
This work was funded by the UK Medical Research Council (Grant #: MR/T04487X/1). We thank participants at the Indonesian Health Economics Association Conference (InAHEA) 2021 for their invaluable feedback.

to the concept of heterogenous treatment effects: an individual's expected benefit from receiving treatment depending on their characteristics, formalised as the conditional average treatment effect (CATE) function. Proposed approaches for learning optimal treatment assignment rules either try to learn the CATE function in order to assign the policy to those individuals where the estimated CATE shows a benefit of being treated (e.g. Luedtke and Laan (2016)), or they aim to directly estimate the population average benefit resulting from a given policy assignment rule, and search over a class of permissible policies to find the optimal one (e.g. Athey and Wager (2021)). This latter approach may be preferable on the grounds of statistical theory – average treatment effects are easier to learn than the multi-dimensional CATE function – and can lead to more interpretable rules if the class of permissible policies is restricted to simple structures, such as the decision trees proposed by Athey and Wager (2021).

Underlying both the CATE-based and the direct approaches there is a crucial task of estimating so-called nuisance functions: the propensity score, defined as the conditional probability of receiving treatment given covariates, and the outcome regression, defined as the conditional expectation of the outcome given treatment and covariates. These nuisance components need to be estimated well for two reasons: first (assuming no unobserved confounders), to eliminate bias due to confounding, and second, to capture heterogeneity in treatment response. The use of flexible machine learning algorithms to estimate nuisance functions has been increasingly recommended, originally for settings where the interest is in average treatment effects (Van der Laan, Rose, et al. 2011; Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, and Newey 2017), and more recently for settings that study the CATE (Künzel et al. 2019; Athey, Tibshirani, et al. 2019). However, the choice of estimators for nuisance components has received less attention in the optimal policy learning problem. One particularly challenging setting common in evaluating health policies is that of rare outcomes.

In settings with rare outcomes - infant mortality, or incidence of side effects from a new medication, for example - little information is contributed by the majority of observations, resulting in high variance of the estimators of treatment effects (Franklin et al. 2017).[1] For the purpose of optimal policy estimation, the concern is that this lack of information may prevent good estimation of nuisance components and the CATE function. It is therefore crucial to

[1]Franklin et al (2017) compared propensity score methods for estimating treatment effects on rare outcomes and find regression adjustment and inverse propensity weighting (IPW) produce lower bias and MSE for binary outcomes

understand any resulting spillover effects of this variance on optimal policy estimation. One specific ML algorithm in particular, Bayesian Additive Regression Trees (BART), has been found to exhibit good performance in estimating outcome regression functions in settings with rare outcomes and heterogeneous treatment effects (Hu, Ji, et al. 2021).[2] The good performance of BART for estimating causal parameters has been further demonstrated in a high profile data challenge (Carnegie et al. 2019). However, to our knowledge, BARTs have yet to be considered for estimating optimal policy allocation rules.

Policy learning from observational data has a long history in the statistics literature (Manski 2004; Qian and Murphy 2011; Luedtke and Van Der Laan 2016). Our approach more closely mirrors the machine learning strand of the policy learning literature (Bertsimas and Kallus 2020; Kallus and Zhou 2021), which relies on data-driven approaches, and relates the loss function to some practical criteria, such as utility. The objective of this paper is therefore to, through a case study and simulations, compare a range of state-of-the-art machine learning approaches that have been specifically designed to estimate either average or heterogeneous (conditional) treatment effects, then use these components as inputs to the tree-based optimal policy learning algorithm proposed by Athey and Wager (2021). Our simulation study creates synthetic datasets with known response surfaces for the potential outcomes, resulting in known CATE functions and oracle policy assignment rules, which we can use to evaluate performance.

To concretely motivate this work, we first examine a commonly studied rare outcome in the health economics literature - infant mortality - which has been shown to exhibit significant heterogeneity in Low and Middle Income Countries (LMICs). In particular, Kreif et al. (2021) show evidence of a reduction in infant mortality resulting from the expansion of social health insurance in Indonesia; however, this reduction was only statistically significant among those who were beneficiaries of a contributory health insurance scheme, and not among those who were recipients of subsidised insurance for the poor. For the contributory insurance group, a Causal Forests approach uncovered significant heterogeneity in the CATEs for the main mechanism behind infant mortality reduction, birth attended by a health professional. Due to the rare nature of the infant mortality outcome, direct estimation of CATEs was not attempted in the previous analysis.

[2]Hu and Gu (2021) study the causal effects of multiple treatments on rare outcomes, comparing Bayesian Additive Regression Trees (BART), regression adjustment on multivariate spline of generalized propensity scores (RAMS), and IPW, using simulations and a case study (Hu and Gu 2021)

Our simulation scenarios cover common settings of different sample sizes, varying degrees of heterogeneity, and treatment overlap. We evaluate the methods by calculating Root Mean Squared Errors (RMSE) and Mean Absolute Errors (MAE) for CATEs, ATEs, and the advantage of deploying each learned policy (see Section 2). Although the methods we consider perform similarly in settings with normal outcome prevalence, we find that their accuracy diverges significantly in settings with rare outcomes - especially as heterogeneity in the CATE increases. In small samples, Bayesian additive regression trees often obtain better estimates than other methods studied in this paper. The NDR-learner also performs well as sample size increases. These results have clear implications for both researchers and policy-makers, who may need to carefully consider choice of methodology when learning optimal policy rules in settings with rare outcomes.

2. The optimal policy estimation problem

2.1. **Notation and setup.** We build on Athey and Wager (2021). We have access to observational data, in the form of independent and identically distributed samples (X_i, W_i, Y_i), where X_i is a vector of individual covariates, $W_i \in \{0, 1\}$ is a binary treatment and $Y_i \in \{0, 1\}$ is the (binary) outcome of interest. Individual level causal effects are defined as $\tau_i = Y_i(1) - Y_i(0)$, where $Y_i(w)$ denotes the potential outcome if the treatment had been set to $W_i = w$ (Rubin 1974). We want to use this data to learn a policy allocation rule π which maps $X_i \in \chi$ into a binary treatment decision, ie. $\pi : \chi \to \{0, 1\}$, for policies in a pre-specified policy class Π.

The traditional causal target parameter when evaluating a binary treatment is the difference between expected potential outcomes when everyone in the population is treated compared to when no one is treated, $\Theta = E[Y_i(1) - Y_i(0)]$ - in other words, the average treatment effect (ATE). We can think of the ATE as the expected benefit of deploying a very simple policy, where irrespective of covariates, everyone is assigned to treatment, and this policy is compared to when no one receives treatment. Generalising the ATE for an arbitary policy π, we define the utility of deploying a policy π relative to treating no one as

$$V(\pi) = E[Y_i(\pi(X_i)) - Y_i(0)]. \tag{1}$$

Intuitively, the optimal policy π_* is the one (within the policy class) that maximizes the quantity $V(\pi)$.

We also define the utilitarian regret (Manski 2009) from deploying a policy π relative to the best policy in the class Π as

$$R(\pi) = \max\{V(\pi') : \pi' \in \Pi\} - V(\pi). \tag{2}$$

The formal goal of policy learning is to derive a policy $\hat{\pi} \in \Pi$ from observational data, with the guarantee that the regret $R(\hat{\pi}) = O_P(1/n)$. To achieve this, further assumptions need to be made on the data generating process, and the complexity of the policy class needs to be controlled.

Here, we make the usual (strong) assumption of no unobserved confounding, such as $\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp W_i|X_i$, and overlap, $0 < e(X_i) < 1$, where $e(X_i) = Pr(W_i = 1|X_i)$ is the probability of receiving the treatment, given observed covariates (the propensity score).

Another building block of the optimal policy estimation problem is the conditional average treatment effect function (CATE), defined as the expected difference between the potential outcomes as a function of the covariate profile:

$$\tau(x) = E[Y_i(1) - Y_i(0)|X_i = x]. \tag{3}$$

Note that the ATE is the expectation of the CATE function over the empirical distribution of the covariates, $\Theta = E[\tau(x)]$.

It can be shown (Athey and Wager 2021) that the value of a policy $V(\pi)$ can be written in terms of the CATE as $V(\pi) = E[\pi(X_i)\tau(X_i)]$, and from this the so-called policy advantage, $A(\pi)$ can be defined as:

$$A(\pi) = 2V(\pi) - E[\tau(X_i)]. \tag{4}$$

$A(\pi)$ can be interpreted as an improvement achieved by the policy π over a random treatment assignment baseline. This is the quantity that will be the target of the optimisation.

2.2. **The double-robust score.** The main assumption made by Athey and Wager (2021) is that we have access to a double-robust estimator for the ATE, a so-called double robust score, where the estimator is formed by taking the average: $\hat{\Theta} = \frac{1}{n}\sum_{i=1}^{n} \hat{\Gamma}_i$. The double-robust score is equivalent to the score used in the double-machine learning estimator (Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, and Newey 2017) based on the the augmented inverse probability of treatment weighted estimator for the ATE (J. M. Robins and Rotnitzky 1995) defined as

$$\hat{\Gamma}_i = \hat{m}(X_i, 1) - \hat{m}(X_i, 0) + \frac{W_i - \hat{e}(X_i)}{\hat{e}(X_i)(1 - \hat{e}(X_i)}(Y_i - \hat{m}(X_i, W_i)), \tag{5}$$

where $m(x, w)$ is the counterfactual response surface defined as $E[Y_i(w)|X_i = x]$ which under unconfoundedness can be identified as $E[Y_i|X_i, W_i]$, and its estimate we denote as $\hat{m}(X_i, W_i)$. Constrution of the DR score requires three nuisance parameter estimates: the expected potential outcome under control $\hat{m}(X_i, 0)$, the expected potential outcome under treatment $\hat{m}(X_i, 1)$, and the expected outcome under the treatment actually received $\hat{m}(X_i, W_i)$. Note that the first part of the formula, $\hat{\tau}_m = \hat{m}(X_i, 1) - \hat{m}(X_i, 0)$ is a non-double robust estimate of the CATE, while the second part is a double-robust adjustment term, where the prediction error of the observed outcome is inverse weighted with the estimated propensity score $\hat{e}(X_i)$.

Athey and Wager (2021) show that to achieve the regret guarantees, the components of $\hat{\Gamma}_i$ - the so-called nuisance functions - need to be estimated via machine learning. In this paper we considers several ways to estimate the nuisance functions in this double robust score, motivated by our setting of rare outcomes. We detail the estimators considered in Section 3. Athey and Wager (2021) consider a somewhat modified double-robust score that directly uses estimates of the CATE obtained from the Causal Forests method, and plugs this into a double-robust score, using a different nuisance function $\hat{m}(X_i)$ the expectation of the outcome conditional on covariates but marginalized over the treatment groups. We also consider this approach, and describe the modified score in Section 3.

2.3. **The optimisation.** Regardless of the way $\hat{\Gamma}_i$ is estimated, in the following step it is used to construct the final target of optimisation $\hat{A}(\pi)$, as

$$\hat{A}(\pi) = \frac{1}{n}\sum_{i=1}^{n}(2\pi(X_i) - 1)\hat{\Gamma}_i \tag{6}$$

Intuitively, we now require an algorithm that can search the space of permissible policy rules and find the rule π that the achieves the largest value of $\hat{A}(\pi)$. In the case where the outcome is harmful - in our setting, the outcome of interest is infant mortality - we simply change the sign of $\hat{A}(\pi)$.

(Athey and Wager 2021) propose using a depth-k decision tree, found via exhaustive tree search, to find the optimal policy within the class of depth k trees[3]. We also consider a "plug in" policy allocation rule based on the estimated CATE, where all individuals with a negative estimated CATE ($\hat{\tau} < 0$) receive treatment. In this case, there is no restriction on the policy class, but if the $\frac{1}{\sqrt{n}}$-rate estimation of the CATE is not possible, treatment assignment rules derived from a simple sign rule may not be asymptotically minimax-optimal (Hirano and Porter 2009).

To summarise, the policy estimation process in this paper consists of four steps.

1. We estimate the nuisance parameters: $\hat{m}(X_i, W_i = w)$ and $\hat{m}(X_i)$.

2. Using the nuisance parameters, we estimate CATEs (for all methods except Double-Debiased ML (Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, and Newey 2017))

3. Using the nuisance parameters and the CATEs (if necessary and available) we construct the DR-scores for each method, $\hat{\Gamma}_i$.

4. We estimate the optimal policy assignment using exhaustive tree search, searching for a treatment rule that maximises the objective $-\hat{A}(\pi)$.

5. For each method we report the estimated value of $-\hat{A}(\pi)$ under the selected tree-based policy, the CATE based plug-in policy. We contrast these to simple plug-in policies obtained directly using the CATE, as well as a simple policy where we treat everyone. We also report the estimated ATE as the mean of each set of scores(Chernozhukov, Demirer, et al. 2018).

3. Machine Learning estimation of double-robust scores

In this section we review the approaches we selected to estimate the DR scores used for optimal policy estimation. Each ML algorithm will yield nuisance functions $\hat{m}(W_i, X_i)$ and $\hat{e}(X_i)$, while the Causal Forests approach also uses the nuisance function $\hat{m}(X_i)$.

3.1. **Bayesian Additive Regression Trees (BART).**

Bayesian Additive Regression Trees (BART) is a machine learning method known to model response surfaces very well in situations with small effect sizes (Hahn et al. 2020), but to date it has not been studied in the double-robust policy learning setting. BART , a "sum-of-trees" model constrained by an influential prior, was extended to binary classification by Chipman et al. (2010) and adapted to the causal inference setting by J. L. Hill (2011). Chipman directly models the response surface for a binary outcome in a probit framework:

[3]We use the `policytree` package in `R` to obtain the policies.

$$m(w, X_i) = E[Y_i|W_i = w, X_i] = \Phi\{\sum_{j=1}^{J} g_j(w, X_i; T_j, A_j)\}, \tag{7}$$

where ϕ is the c.d.f. of the standard normal distribution, T_j denotes a single regression tree and A_j is the set of its associated parameters, and $g_j(w, X_i; T_j, A_j)$ represents the conditional mean assigned to the particular node associated with covariate profile X_i and treatment w in the jth regression tree. The imposed prior regularizes the fit of each regression tree by keeping individual effects small (Chipman et al. 2010), with each tree explaining only a portion of the response surface. Iterations of a Bayesian backfitted MCMC construct and fit separate residuals, which are effectively S MCMC samples from an induced posterior distribution. To perform causal inference, missing potential outcomes are predicted for each observation using imputed c.d.fs, and treatment status w using each sample draw S (J. L. Hill 2011; J. Hill et al. 2020). CATEs are then easily obtained from $\sum_{s=1}^{S} m(1, X_i) - m(0, X_i)$. We create the double-robust score as in Equation 5 using BART estimates of $\hat{e}(X_i)$ and the conditional response surfaces.

BART's desirable properties include its ability to handle non-linearities and multi-way interactions between covariates without researcher input (Tan and Roy 2019). Further, because the algorithm yields posterior samples, point and interval estimates are easily obtained by taking averages or caculating quantiles of resulting draws. In our setting, casual BART is modeled with the estimated propensity score $\hat{e}(X_i)$ (also obtained using BART) included as a covariate, as this has been shown to improve performance (Hahn et al. 2020). Note that the original BART algorithm is not prevented from extrapolating over areas of the covariate space which fall outside of common support (Hu, Gu, et al. 2020; Hu, Ji, et al. 2021). Hu et al (2020) propose discarding units with a large variability in predicting potential outcomes to address this issue. In the simulations, we could potentially see improvement in the BART estimates in larger samples by adopting this procedure, as we have artificially constructed scenarios with poor overlap, however in this setting we do not use a discard rule. We perform 1200 MCMC draws, with the first 200 treated as burn-in.

3.2. **Model Agnostic Double-Robust Machine Learning.** We next consider Double Machine Learning approaches, originally proposed by Chernozhukov and colleagues (Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, and Newey 2017; Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and J. Robins 2018) to estimate average treatment effects.

The general idea is that estimators of ATEs[4] can be constructed as double-robust scores - which we will use in the optimal policy estimation - using nuisance components estimated by machine learning. The approach is referred to as "model agnostic" in that it can be implemented using a variety of supervised machine learning methods. For consistency, across all the model-agnostic methods we estimate the nuisance functions using regression forests.

3.2.1. *Generic Double-Debiased ML (DML).* We first consider the score from the the double-debiased ML approach developed by Chernozhukov and colleagues (Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, and Newey 2017; Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and J. Robins 2018) for estimating ATEs. Using results from semiparametric statistical theory on asymptotically linear estimators, the authors derive a double-robust score (see Equation 5), where the nuisance components - propensity scores and outcome regression functions - are estimated using ML algorithms. If these ML estimators converge fast enough to the truth, and if a so-called cross fitting procedure, detailed below, is used - the resulting estimator for the ATE is asymptotically linear and consistent.

As proposed by the authors, we estimate the scores using cross fitting over K folds: regression forests with 2000 trees[5] are trained on $K-1$ folds of the data then nuisance parameters are predicted using the forest fits for the kth holdout fifth fold, with the process then repeated K times to obtain a score for each observation. In our implementation, we use $K=5$ folds.

Note that Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and J. Robins (2018) do not advocate directly estimating CATEs from the scores obtained via DML. Instead, they suggest linearly regressing the obtained scores on covariates (e.g. indicators of subgroups of interest) to learn about the relative importance of various variables in predicting the CATEs. They call these coefficients from the linear model the best linear predictors (BLP of CATEs). Therefore, for our purposes of optimal policy rule evaluation, we compute the double-robust scores only, and do not obtain CATEs from this method.

3.2.2. *DR-Learner.* Kennedy (2020) extend the DML framework to obtain estimates of the CATE, by implementing a two-step procedure. First, double-robust scores (as in Equation 5), referred to as "psuedo-outcomes", are constructed, then these are regressed on the covariates X, using supervised ML. The procedure is performed using a sample splitting procedure similar to the method described above, with some slight differences. First, $\hat{\mu}(w, X_i)$ and

[4]And, in some settings, CATEs

[5]We use the default parameters in the `grf R` package.

$\hat{m}_w(X_i)$ are each obtained from two separate training samples, and are used to construct the DR score (this is in contrast to the DML method, which estimates both nuisance functions using the same subsample). Second, these separate nuisance estimates are used to predict the scores for a third, final sample, which is then regressed on covariates within that third sample to obtain the CATE, ie, $\hat{\tau}_{dr}(Z) = E_n[\hat{\Gamma}(Z)|X = x]$. In contrast to the DML method above, the only requirement for valid estimation of the CATE is that the second-stage regression satisfies mild stability assumptions and its squared-error loss functions are known - therefore a more direct estimate of the CATE (rather than its Best Linear Predictor) is assumed possible. This is repeated three times, where each sample is used once in the final step, and the average of the estimators is used as the final estimate of τ.

For our purposes of optimal policy estimation, we construct the scores exactly as with the DML procedure described above (Equation 5), however these are taken as averages across the repetitions used in out-of-sample prediction of the CATEs. We use regression forests to obtain nuisances, and the default 5 folds in the `causalDML R` package.

3.2.3. *Normalized-DR Learner.* Knaus (2020) proposes the NDR learner for CATE estimation, which differs from the DR learner described above in that it normalizes the inverse probability weights used in the DR score to help stabilize individual treatment effect estimates. Knaus observes that point estimates in the DR-learner can be expressed as $\hat{\tau}_{dr}(Z) = \sum_{i=1}^{N} \alpha_i(\hat{\Gamma}_i$, when α_i can be calculated and is the weight that each observation receives in the second-stage regression above. The proposed NDR-learner procedure ensures that weights on the individual outcome residual portion of the $\hat{\Gamma}$s, $Y_i - \mu(w, X_i)$, sum to one. Note that this method restricts the underlying ML methods to linear smoothers only.[6]

The procedure follows that of the DR-learner above: first, it construct the same scores as the DR-learner, and performs the normalization step is after the score is regressed on X_i (after the final step in the previous subsection). The CATEs estimated by the method have been found to perform well in simulations (Knaus 2020). Because the scores are identical to those of the DR-learner, we report policy advantage estimates for the NDR-learner only. We compare CATE estimates between the DR and NDR-learner in Section 6.1.

[6]A linear smoother is an operation by which transformed variables can be expressed as a linear transformation of observed values. That is, $\hat{X} = (\hat{x_1}...\hat{x_n})^t$ can be written in the form $\hat{X} = SX$, where the smoother matrix S does not depend on the original X (Buja et al. 1989). A simple example of a linear smoother is a moving average. ML examples include tree-based algorithms or ridge regression.

3.3. **Causal Forests.** Finally, we turn our attention to the method of generalised random forests (Wager and Athey 2018; Athey, Tibshirani, et al. 2019), and its implementation for heterogeneous treatment effects estimation, Causal Forests Athey and Wager (2019). Causal forests are not model agnostic learners, i.e. they rely on a specific machine learning method - a version of random forests - to both estimate CATEs and obtain the double-robust scores. We consider two implementations of the Causal forests approach to obtaining nuisance functions and CATEs used in constructing the double-robust scores or optimal policy estimation: the so-called Honest Causal Forests, and a novel extension we propose, cross-fitted Causal Forest.

3.3.1. *Honest Causal Forests.* In brief, the Causal Forests estimator relies on modified regression forests - generalised random forests - to find small neighborhoods (leaves of a tree) where the CATEs is constant, by regressing the residualised outcomes on the residualised treatment variable,[7] and partitioning the data into leaves to maximise the between-leaf heterogeneity in the estimated treatment effects, defining so-called Causal Trees. To reduce noise stemming from using individual trees, this procedure is done many times on bootstrap samples, forming a Causal Forest. The causal forests are then used to calculate $\alpha_i(x)$ weights for each observation, based on how frequently an observation was used to estimate the treatment effect at x. The estimator for the CATE is then constructed using these weights and the nuisance components as

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} \alpha_i(x)(W_i - \hat{e}(X_i))(Y_i - \hat{m}(X_i))}{\sum_{i=1}^{n} \alpha_i(x)(W_i - \hat{e}(X_i))^2} \tag{8}$$

Wager and Athey (2018) call their forests "honest" in that each observation is only used to estimate within-leaf treatment effects $\hat{\tau}_i$, or to decide where to place splits within a tree, but not both.

Although the double-robust score resulting from a causal forest is conceptually identical to Equation 5, the ability of the causal forest to directly estimate CATEs allows for the direct use of $\hat{\tau}(X_i))$

$$\hat{\Gamma}_i = \hat{\tau}(X_i) + \frac{W_i - \hat{e}(X_i)}{\hat{e}(X_i)(1 - \hat{e}(X_i)}(Y_i - \hat{f}(X_i) - (W_i - \hat{e}(X_i))\hat{\tau}(X_i)), \tag{9}$$

[7] Residualisation is performed by subtracting the predicted outcome $\hat{m}(X_i, W_i)$ and and the estimated propensity score $\hat{e}(X_i)$ from the outcome and treatment indicators, respectively.

where $\hat{f}(X_i) = E[Y_i|X_i = x]$.

We use the `grf R` package to estimate and fit the causal forests, tuning all parameters and growing 2,000 trees.[8]

3.3.2. *Cross-Fitted Causal Forests (CFTT).* In this novel implementation, we modify the causal forest to cross-fit out-of-sample estimates on K randomly generated folds of the data. First, we fit a causal forest based on $K-1$ folds. We then use these fitted forests to predict the nuisance functions and CATEs on the kth holdout fold. DR scores are constructed from these estimates as in Equation 9. The procedure is performed K times to obtain predictions for each observation, where each kth fold acts as the holdout fold once. The entire process is then repeated t times (i.e., the sample is re-split into K' new folds and forests are predicted for each), and results across each repetition are averaged to obtain final estimates.

Although the Honest Causal Forest uses out-of-bag sampling when estimating CATEs (i.e., an observation may not be used to estimate τ if it has been used to determine a split in that particular regression tree), over the space of the entire causal forest, it could still be the case that overfitting is occurring if the same observations are being used for obtaining both nuisance functions and the CATE prediction across many trees. By restricting/splitting the sample before the tree bootstrapping occurs, there is a reduced likelihood for the same observations being used to determine splits (and therefore not used in the within-leaf predictions). Our cross-fitting procedure using testing and training (hold-out) data is standard in many ML applications and modifies the original Honest Forest algorithm only slightly. Within the individual fold fitting and prediction, we maintain tree honesty (so the same observations *within the training fold* that are used to obtain tree splits are not used to obtain CATEs). We use $K = 5$ and repeat the procedure $t = 4$ times in our implementation.[9] As above, we tune all parameters and fit 2,000 trees for each fold.

4. Motivating Case Study: Indonesian National Health Insurance Programme

Building on Kreif et al. (2021) we aim to explore heterogeneity in the effect of subsidised health insurance on the reduction in infant mortality, but instead of focussing on the estimation of the individual level CATEs, we explore heterogeneity in the treatment effect across

[8]See Tibshirani et al. (2021) for further details on Causal Forest tuning parameters. We grow 100 trees to tune parameters, and repeat the tuning process 500 times.

[9]Code is available at `github.com/jhatamyar/OptPolicyRare`

certain groups. Then we turn to the estimation of optimal policy assignment rules, using each of the methods described in Section 3.

4.1. **Data.** The dataset consists of births between 2002 and 2014, extracted from the Indonesian Family Life Survey (IFLS), a longitudinal household survey (Strauss et al. 2016). The unit of observation is a birth, while the treatment is defined as for a given birth, whether a mother was covered by subsidised health insurance in the year of the birth. The control group consists of those births where no insurance was reported in the year of the birth. The outcome of interest is infant mortality, measured as child death before the first year of life.

To deal with the challenge of self-selection into health insurance, Kreif et al exploits variation in the expansion of both contributory and subsidized health insurance schemes, across provinces and over time. The observed confounders take into account information on known predictors of infant mortality, as well as the eligibility criteria of subdsidised health insurance. They include the mother's characteristics (age, education, wealth in quintiles) and household characteristics (social assistance, experienced a natural disaster, rurality, availability of health services: a village midwife, birth clinic, hospital). Kreif et al also control for region effects that capture unobserved confounding factors that are common within regions and time-invariant. As reported by Kreif et al. (2021), births under subsidised insurance were more likely to be from a rural household and from mothers who are older at birth, less likely to have studied at university and more likely to have only elementary school education, belong to lower wealth quintiles, and receive social assistance programmes, compared to those with no insurance. Hence, good estimation of the nuisance functions is crucial to reduce confounding in this study. As demonstrated by Kreif et al. (2021), inverse probability weighting using a logistic regression based propensity score was effective in creating balance across the observed covariates.[10]

4.2. **Results of the case study.** We report estimates of the average treatment effect of subsidized health insurance on infant mortality, derived from the double robust scores obtained using each method, and their resulting policy advantage (Equation 6) using using depth 2 trees in Table 1. There is slight disagreement across the methods in terms of ATE estimates - Causal Forests are found to estimate the largest (most negative) average treatment effect, while the NDR-learner and BART result in smaller estimates. There is also

[10]Balance checks using the ML based propensity scores will be available soon

variation in the calculated policy advantage A_i of a depth-two policy tree using scores obtained by each method. As this variation exhibits a difference of nearly half a percentage point (-0.023 for DML vs -0.019 for CFTT and BART) across estimated advantages, it is unclear which ML method is most trustworthy.

TABLE 1. ATE: Insurance and Infant Mortality, IFLS

	Estimated ATE		**Policy Advantage**	
	ATE	(SE)	Ai	(SE)
Kreif et al 2021	-0.005	(0.005)		
DML	-0.004	(0.005)	-0.023	(0.005)
NDR	-0.003	(0.006)	-0.021	(0.005)
CF	-0.005	(0.005)	-0.020	(0.005)
CFTT	-0.004	(0.005)	-0.019	(0.005)
BART	-0.003	(0.005)	-0.019	(0.005)

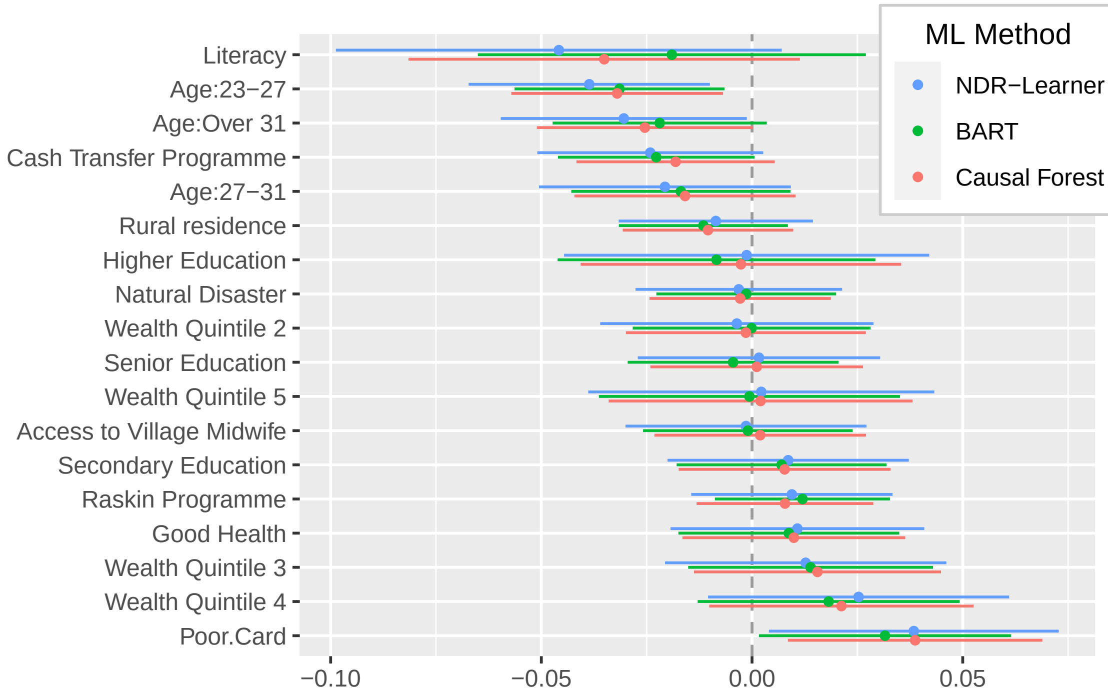
This table reports estimated ATE of subsidized health insurance on infant mortality using the IFLS data, i.e. the mean of double-robust scores, for the various ML methods in this paper, as well as the estimated Policy Advantage $\hat{A}_i(\pi)$ (Equation 6) for each method. We also report the Causal Forest estimates of ATE from Kreif et al (2021) for comparison.

Figure 1 depicts the Best Linear Predictor of the Group Average Treatment Effects (Chernozhukov, Demirer, et al. 2018) for a selected group of covariates, using scores from an Honest Causal Forest (CF). The target parameter is $E[\hat{\tau}(X_i)|G]$, where G indicates membership in some subgroup. Practically, this is obtained by linearly regressing the estimated double-robust scores on the covariates G of interest. While we find some evidence of heterogeneity, across the groups, the confidence intervals around the group average treatment effects overlap, with one group (those in possession of a card showing poverty status) that surprisingly seems to be harmed by the health insurance programme.

Table 2 shows the estimated advantage of learned depth-two policy trees compared to a plug-in rule where treatment is assigned to anyone with a negative CATE, and a policy where all individuals are treated. We compare learned policy trees which use all covariates in X to a tree restricted to using only those covariates studied in our BLP of GATE analysis (Figure 1).[11] It also reports the percent of the population treated under the respective policies.

[11] Although the scores used to obtain the policy trees are equivalent for both of these settings, the tree is restricted to a subset of covariates in the second version. This demonstrates the practical use of decision-rule

FIGURE 1. Best Linear Predictor of Group ATEs: Infant Mortality



This figure depicts the results of a linear regression of estimated double-robust scores on a selection of covariates from the IFLS data.

Although the estimated advantage $\hat{A}(\pi)$ differs between ML methods, the actual learned policy (and resulting percent of individuals treated) is identical across methods when all covariates are allowed to be used in the decision tree. When the class is limited to a subset of covariates (whether for interpretability or equity reasons), the policy advantage $\hat{A}(\pi)$ is reduced - yet still results in a greater potential reduction in infant mortality than the plug-in policy. For a plug-in policy which treats all individuals with an estimated $\hat{\tau} < 0$, the portion of treated observations varies along with the estimated policy advantage, which is lower than the estimated advantage for a depth-two tree based policy regardless of covariates or ML method used. The plug-in policy is estimated to result in a greater reduction in mortality compared to one which treats 100% of the sample (last column).

---

based policy classes: a policy maker may not want their treatment assignment to be allowed to rely on certain covariates (such as race, gender, or location) for equity reasons.

Table 2. Policy Advantages, Subsidized Health Insurance

Policy Rule:	All $X_i(tree)$		Limited $X_i(tree)$		Plug-in ($\hat{\tau} < 0$)		$\hat{e}_i = 1$
	$\hat{A}_i(\pi)$	% TX	$\hat{A}_i(\pi)$	% TX	$\hat{A}_i(\pi)$	% TX	Ai
DML	-0.023	0.825	-0.017	0.741	-	-	-0.004
NDR	-0.021	0.825	-0.017	0.741	-0.008	0.680	-0.003
CF	-0.020	0.825	-0.017	0.741	-0.010	0.887	-0.005
CFTT	-0.019	0.825	-0.017	0.741	-0.009	0.725	-0.004
BART	-0.019	0.825	-0.017	0.743	-0.001	0.641	-0.006

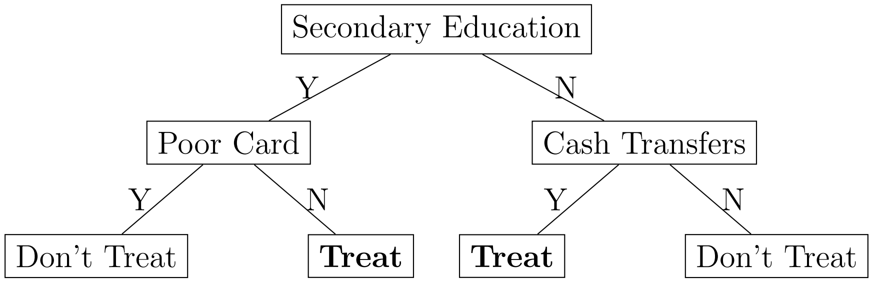
This table reports estimated policy advantages for a variety of policy types, as well as the percent of individuals in the sample who would be treated under each policy. The first two columns show the estimated $\hat{A}_i(\pi)$ for a depth-two decision tree which allows for treatment criteria to be based on all available covariates. The second two columns limit the tree decision criteria to those covariates examined in Figure 1. The plug-in policy uses estimated CATEs from each method to assign treatment based on a rule where $\hat{\tau} < 0$. No policy is learned using DML for this class, as this method does not allow for estimation of CATEs. The final column depicts the advantage of a policy which treats the entire sample - note that in most cases, this is identical to the ATE in Table 1.

Figure 2 depicts the treatment assignment rule from a depth-two tree learned from the BART double-robust scores. This tree is restricted to the subset of covariates used in the BLP of GATE analysis. It is clear that the exhaustive tree-search algorithm (and the scores which this algorithm utilizes) picks up on the undesirable positive effect of the Poor Card, as it yields a treatment assignment rule which does not treat individuals with a Poor Card if they have secondary education.[12] The policy tree also indicates the potential importance of secondary education and cash transfers as drivers of treatment effect heterogeneity.

Our case study also demonstrates the potential downside of a plug-in policy which simply treats those with a negative estimated CATE. Depending on the method used to obtain the CATEs, the distribution of the estimates may vary substantially, as is seen in Figure 3, which contrasts the the BART and Causal Forest methods. Therefore, a the efficacy of a policy using the plug-in treatment rule may be particularly reliant on the accuracy of the CATE estimates.

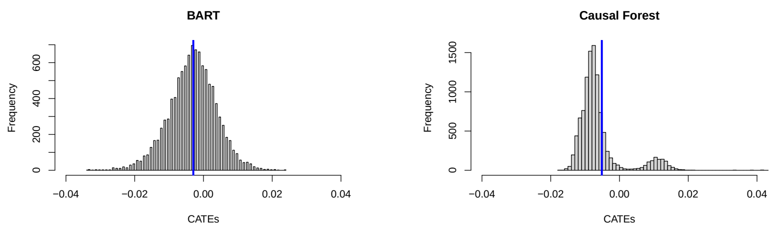
[12]It is likely the case that a policy maker would not wish to exclude individuals from treatment should they be in possession of a card indicating poverty status. The learned policy would be easily modified by dropping this covariate from the list of acceptable decision criteria in the tree-search algorithm. In this paper, we allow the Poor Card variable to be a treatment decision criteria in order to demonstrate the ability of the policy tree class to pick up on potential drivers of heterogeneity.

FIGURE 2. Depth-Two Learned Policy: BART



A simple depth-two decision tree for assigning treatment, learned from double-robust scores obtained via BART.

FIGURE 3. Distribution of CATEs



(A) (B)

The left panel depicts the distribution of estimated CATEs of subsidized insurance on infant mortality using BART. The right panel depicts the estimates obtained using Causal Forests.

5. SIMULATIONS

5.1. **Simulation Design.** In all scenarios, we consider three sample sizes of $n = 500$, $n = 1000$ and $n = 2000$. We generate two variations of the binary outcome prevalence: 40%-60% (normal), and 2%-7% (very rare). All scenarios use an approximate 35% treatment prevalence. [13] As in Hu, Ji, et al. 2021, we simulate 10 covariates with continuous $X_1, ..., X_5$ drawn from the standard normal distribution and categorical $X_6, ..., X_{10}$ from Bernoulli(0.5).

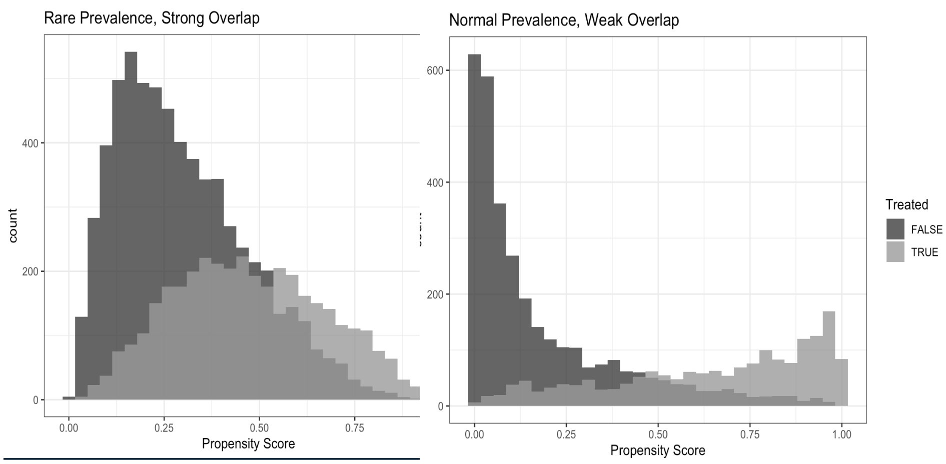
We induce confounding by allowing the treatment assignment and response surfaces to both be functions of all the covariates. The true propensity score $e(x)$ is given by:

[13]The goal is to maintain a consistent treatment prevalence across all scenarios, and comparable outcome prevalence by treatment group across individual scenarios and overlap settings. Details of actual treatment and outcome prevalence for each simulated scenario are shown in the Appendix.

$$\text{logit}(exp(X)) = \alpha + \psi(-0.1X_1 - 0.9X_2 - 0.3X_3 - 0.1X_5 - 0.1X_6 - 0.2X_7 - 0.4X_9 + .5X_{10}) \tag{10}$$

where $\psi \in (1, 2)$ represent moderate and weak overlap respectively, and $\alpha \in (0.8, 1.6)$ chosen to ensure the proportion of treated individuals is around 35 %. Figure 1 depicts the overlap for the strong and weak scenarios.

Figure 4. Moderate and Weak Overlap Settings



(A) (B)

We model response surfaces for the binary outcome drawing from Bernoulli(ρ), where ρ is given by $m(x, w)$ as below. We induce different levels of treatment effect heterogeneity by constructing the following scenarios for response surfaces, which result in a known CATE function and an oracle policy assignment rule for each scenario. We induce various levels of heterogeneity in order to obtain non-trivial treatment assignment rules - if there is no heterogeneity in the CATEs across observations, there is little to gain from a learned policy.

- **Scenario 1:** non-linear effects only in the response surface under treatment.

$$m(X_i, 0) = \text{logit}(exp(\alpha_0 - 0.5X_1 - 0.8X_3 - 1.8X_5 - 0.9X_6 - 0.1X_7)$$
$$m(X_i, 1) = \text{logit}(exp(\alpha_1 + 0.1\text{logit}(X_1) + 0.1\sin(X_3) - 0.1X_5^2 - 0.3X_6 - 0.2X_7)$$

- **Scenario 2:** non-linear effects in both the treated and control response surfaces.

$$m(X_i, 0) = \text{logit}(exp(\alpha_0 + 0.1X_1^2 - 0.2\sin(X_3) + 0.2\text{logit}(X_5) + 0.2X_6 - 0.3X_7)$$
$$m(X_i, 1) = \text{logit}(exp(\alpha_1 + 0.1\text{logit}(X_1) + 0.1\sin(X_3) - 0.1X_5^2 - 0.3X_6 - 0.2X_7)$$

- **Scenario 3:** non-linear effects in both the treated and control response surfaces, but as functions of different covariates.

$$m(X_i, 0) = \text{logit}(exp(\alpha_0 + 0.1X_1^2 - 0.2\sin(X_4) + 0.2\text{logit}(X_5) + 0.2X_6 - 0.3X_8)$$
$$m(X_i, 1) = \text{logit}(exp(\alpha_1 + 0.1\text{logit}(X_2) + 0.1\sin(X_3) - 0.1X_5^2 - 0.3X_7 - 0.2X_9)$$

5.2. **Simulation Procedure.** For each DGP:

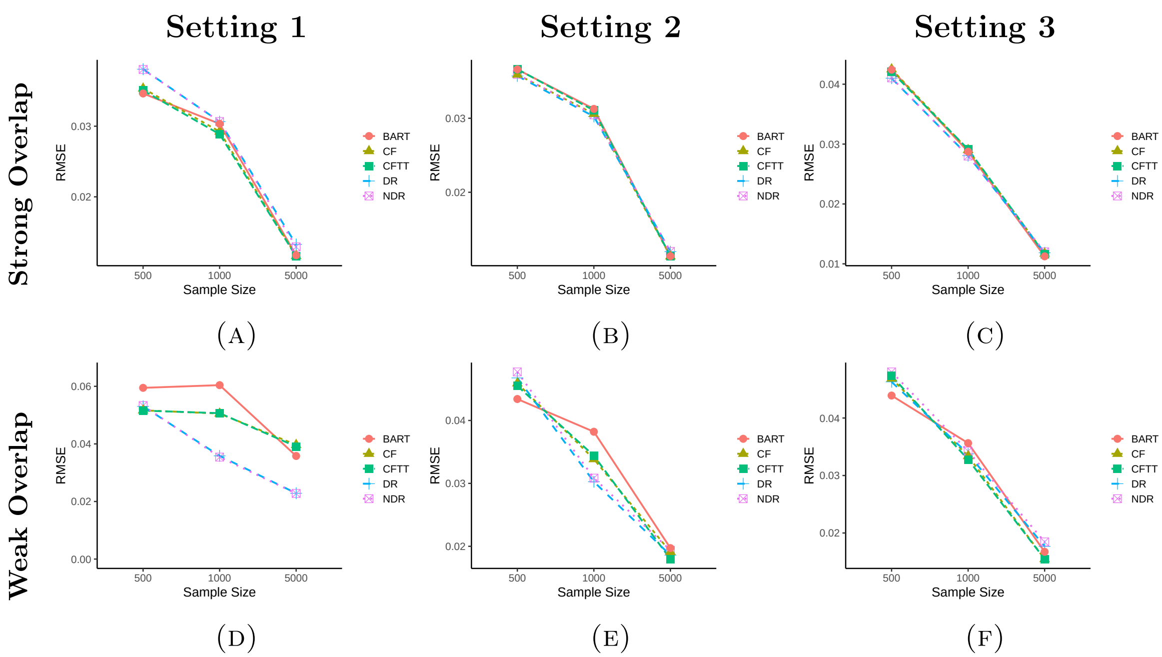
(1) For each repetition:
  (a) Generate data according to scenario and overlap settings
  (b) For each method:
    (i) Estimate nuisance functions
    (ii) Estimate CATEs
    (iii) Construct DR-scores
    (iv) Estimate the optimal 2-level PolicyTree
  (c) Store results
(2) Take averages of estimates
(3) Compute/compare RMSE of CATEs and overall ATE for each method
(4) Compare policy advantages from each method with the "oracle" policy using methods in Athey and Wager (2021), where the oracle policy is derived from a treatment rule which treats all individuals with a *true* negative CATE.

6. Simulation Results

6.1. **Conditional Average Treatment Effects (CATEs).**

Normal Outcome Prevalence We first examine the performance of the various methods when estimating the CATEs. Figure 5 depicts the average RMSE for each method. In settings with strong overlap, the methods perform similarly. However, with weak overlap, the DR and NDR-learners obtain the lowest average RMSE in Setting 1 (Panel (D)). As the degree of heterogeneity increases in Settings 2 and 3, the methods generally perform similarly. While BART shows some advantage in the smallest sample size, as N increases, all other methods outperform BART.

FIGURE 5. RMSE of CATEs: Normal Outcome Prevalence

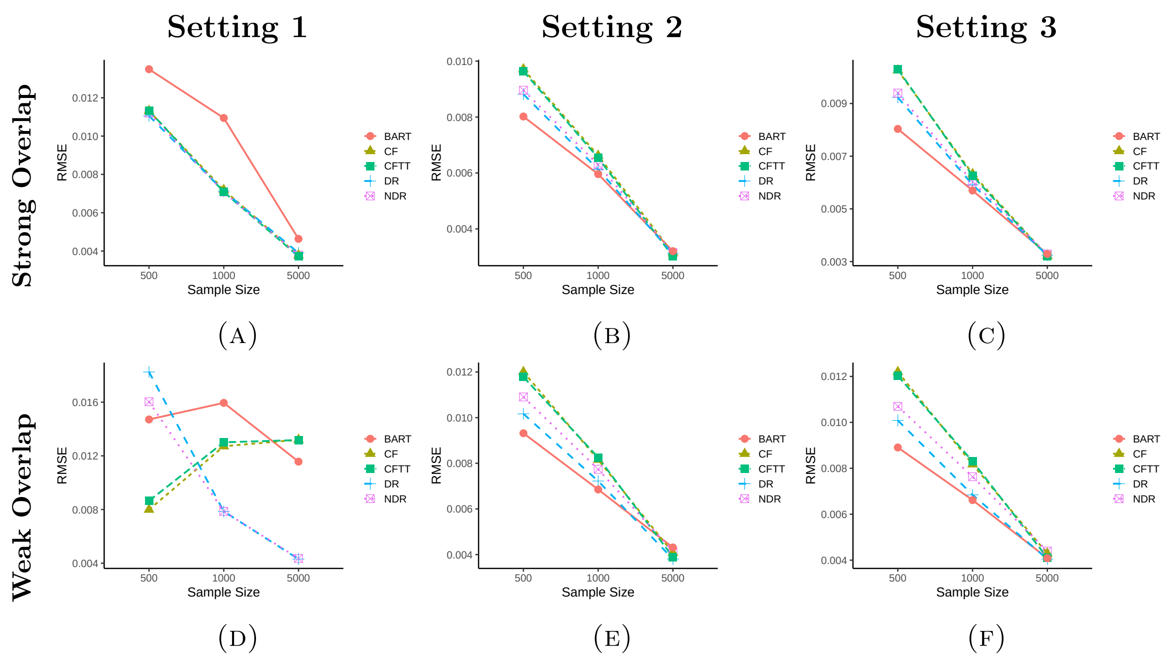


Rare Outcome Prevalence RMSE of the CATE estimates for the rare outcome setting are depicted in Figure 6. We observe that there is more variation across methods in all settings compared to the previous normal outcome prevalence scenario. There are also large differences in performance between Setting 1, the lowest complexity of treatment effect heterogeneity, and Settings 2 and 3. With both strong and weak overlap (Panels (A) and (D)), the DR and NDR learners generally obtain the lowest RMSE of CATE estimates. In the smallest sample size, causal forests perform best in the weak overlap scenario - however, the DR and NDR learners overtake both causal forest and BART as the sample size increases. BART exhibits poor performance in both overlap scenarios for Setting 1. However, as the degree of heterogeneity increases (Settings 2 and 3), BART is a clear winner for obtaining the lowest RMSE of CATE estimates in smaller samples - however, as N increases, the methods again converge to similar performance. RMSE of CATEs are also reported in Table 3.

6.2. **Performance: Average Treatment Effects (ATEs).**

Normal Outcome Prevalence We first examine the RMSE of the overall ATE estimate in the normal outcome setting. Results are depicted in Figure 7. In all settings, BART

Figure 6. RMSE of CATEs: Rare Outcome Prevalence



(A) (B) (C)

(D) (E) (F)

exhibits the worst performance in terms of RMSE for the ATE, with the DR and NDR learner obtaining the lowest RMSE. The DR and NDR learners perform particularly well in settings with weak overlap and larger degrees of heterogeneity (Panels (E) and (F)).

Rare Outcome Prevalence Results for rare outcomes are depicted in Figure 8. Once again, BART exhibits the worst performance in terms of RMSE for the ATE, with the DR and NDR learner obtaining the lowest RMSE in large samples and with greater degrees of heterogeneity. In the case of weak overlap and Setting 1, we see an improvement by the Cross-Fitted Causal Forests over the Honest Causal Forest in small samples (Panel (D)). We depict numeric results for the RMSE of ATE in Table 4.

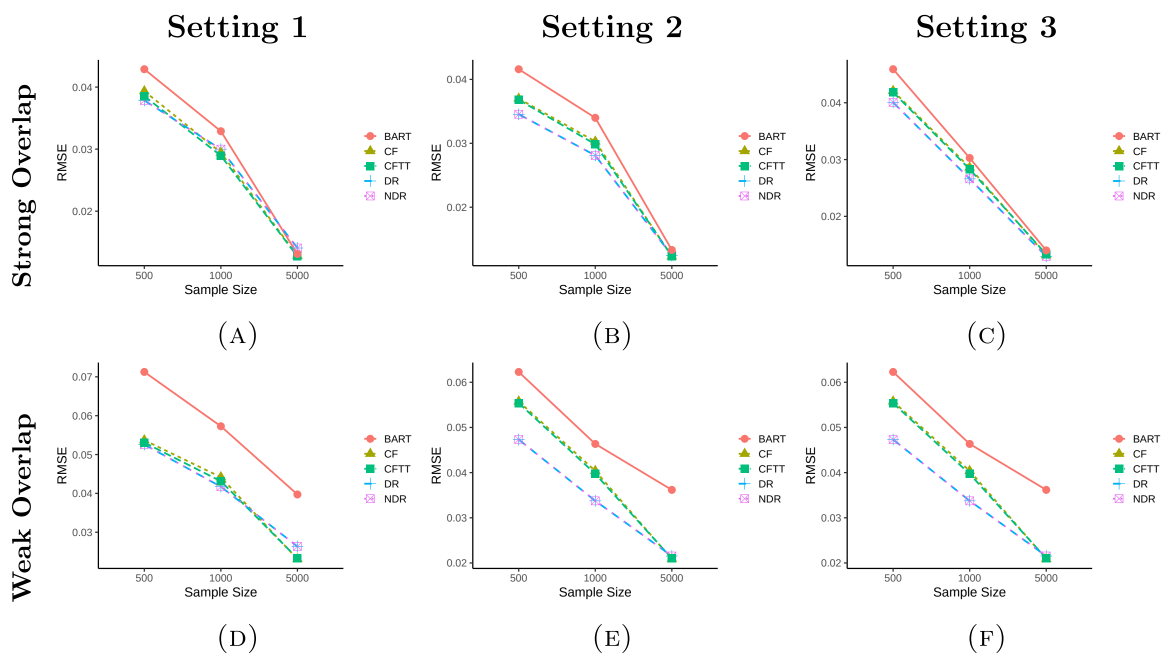
TABLE 3. RMSE of CATEs (SD)

	SETTING 1			SETTING 2			SETTING 3		
Normal: Strong Overlap	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.038	0.031	0.013	0.010	**0.030**	0.012	**0.041**	**0.028**	0.012
	(0.027)	(0.022)	(0.010)	(0.009)	(0.025)	(0.009)	(0.032)	(0.023)	(0.009)
NDR	0.038	0.031	0.013	0.011	**0.030**	0.012	**0.041**	**0.028**	0.012
	(0.026)	(0.022)	(0.010)	(0.010)	(0.025)	(0.009)	(0.033)	(0.023)	(0.009)
CF	**0.035**	**0.029**	**0.012**	0.012	0.031	**0.011**	0.043	0.029	0.012
	(0.027)	(0.020)	(0.009)	(0.010)	(0.025)	(0.009)	(0.032)	(0.023)	(0.009)
CFTT	**0.035**	**0.029**	**0.012**	0.012	0.031	**0.011**	0.042	0.029	0.012
	(0.027)	(0.020)	(0.009)	(0.009)	(0.025)	(0.009)	(0.032)	(0.023)	(0.009)
BART	**0.035**	0.030	**0.012**	**0.009**	0.031	**0.011**	0.042	0.029	**0.011**
	(0.027)	(0.022)	(0.009)	(0.008)	(0.025)	(0.009)	(0.032)	(0.023)	(0.009)
Weak Overlap	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.053	0.036	0.023	0.047	**0.030**	0.019	0.046	0.034	0.018
	(0.038)	(0.028)	(0.015)	(0.037)	(0.024)	(0.013)	(0.035)	(0.029)	(0.013)
NDR	0.053	**0.035**	0.023	0.048	0.031	0.019	0.048	0.034	0.018
	(0.037)	(0.028)	(0.015)	(0.038)	(0.024)	(0.013)	(0.036)	(0.030)	(0.014)
CF	**0.052**	0.051	0.040	0.046	0.034	0.019	0.047	**0.033**	**0.015**
	(0.047)	(0.034)	(0.018)	(0.041)	(0.026)	(0.015)	(0.039)	(0.028)	(0.012)
CFTT	**0.052**	0.051	0.039	0.045	0.034	**0.018**	0.047	**0.033**	**0.015**
	(0.047)	(0.034)	(0.018)	(0.041)	(0.025)	(0.014)	(0.039)	(0.028)	(0.012)
BART	0.059	0.060	**0.036**	**0.043**	0.038	0.020	**0.044**	0.036	0.017
	(0.045)	(0.035)	(0.021)	(0.043)	(0.026)	(0.015)	(0.038)	(0.030)	(0.012)
Rare : Strong Overlap	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.011	0.007	0.004	0.009	**0.006**	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.003)	(0.008)	(0.005)	(0.003)	(0.007)	(0.005)	(0.003)
NDR	0.011	0.007	0.004	0.009	**0.006**	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.003)	(0.008)	(0.005)	(0.003)	(0.008)	(0.005)	(0.003)
CF	0.011	0.007	0.004	0.010	0.007	0.003	0.010	0.006	0.003
	(0.008)	(0.005)	(0.003)	(0.008)	(0.005)	(0.002)	(0.008)	(0.005)	(0.002)
CFTT	0.011	0.007	0.004	0.010	0.007	0.003	0.010	0.006	0.003
	(0.008)	(0.005)	(0.003)	(0.008)	(0.005)	(0.002)	(0.008)	(0.005)	(0.002)
BART	0.013	0.011	0.005	**0.008**	**0.006**	0.003	**0.008**	0.006	0.003
	(0.009)	(0.007)	(0.004)	(0.007)	(0.005)	(0.002)	(0.007)	(0.004)	(0.002)
Weak Overlap	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.018	**0.008**	**0.004**	0.010	**0.007**	0.004	0.010	**0.007**	0.004
	(0.008)	(0.006)	(0.003)	(0.009)	(0.005)	(0.003)	(0.009)	(0.005)	(0.003)
NDR	0.016	**0.008**	**0.004**	0.011	0.008	0.004	0.011	0.008	0.004
	(0.008)	(0.006)	(0.004)	(0.010)	(0.006)	(0.004)	(0.010)	(0.006)	(0.004)
CF	**0.008**	0.013	0.013	0.012	0.008	0.004	0.012	0.008	0.004
	(0.003)	(0.008)	(0.005)	(0.010)	(0.006)	(0.003)	(0.009)	(0.006)	(0.003)
CFTT	0.009	0.013	0.013	0.012	0.008	0.004	0.012	0.008	0.004
	(0.005)	(0.009)	(0.005)	(0.009)	(0.006)	(0.003)	(0.009)	(0.006)	(0.003)
BART	0.015	0.016	0.012	**0.009**	**0.007**	0.004	**0.009**	**0.007**	0.004
	(0.009)	(0.009)	(0.007)	(0.008)	(0.005)	(0.003)	(0.008)	(0.005)	(0.003)

[a] This table reports RMSE of estimated CATEs. The top two panels depict normal outcome prevalence, and the bottom two panels depict rare outcome prevalence.

Based on results for the policy estimation step, which find the optimal policy using double-robust scores, it is surprising to see underperformance of BART for the overall ATE (recall that we estimate the ATE as the average of these scores). We therefore also examine the Mean Absolute Error for the ATE estimation. Results are depicted in Figures 9 and 10. According to this metric, BART performs better in Settings 1 and 2 for smaller samples.

FIGURE 7. RMSE of ATEs: Normal Outcome Prevalence

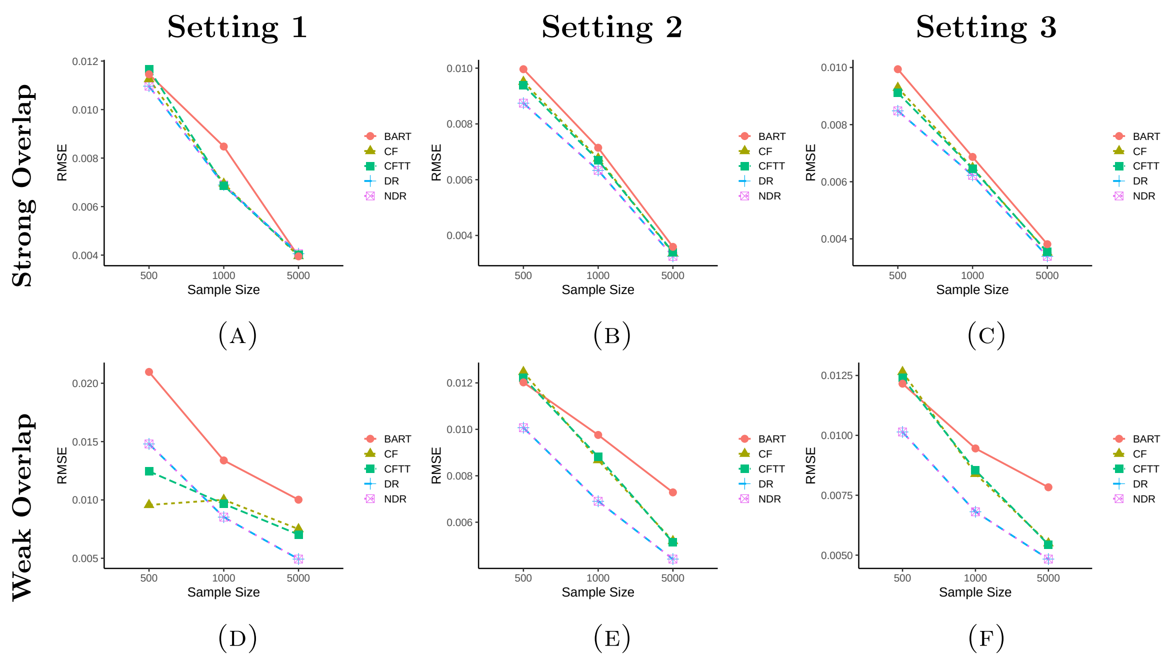


This indicates that the BART scores may, on average, obtain a lower degree of bias which potentially translates into better accuracy in the policy optimization step. We also see the Causal Forest improve using this metric, especially for normal outcome prevalence, but the DR and NDR learners remain the strongest candidates in settings with weak overlap and rare outcomes.

6.3. **Performance: Depth-Two Policy Tree Advantage.** Here we present the main results of the simulations. To facilitate comparison across scenarios, we normalize the RMSE results by the mean oracle policy advantage.[14] Figues 11 and 12 depict results for the normal and rare outcome settings, respectively (numeric results are presented in Tables 5 and 6). Across all settings and scenerios, BART obtains the lowest NRMSE in the policy advangage $\hat{A}_i$ for small settings (N = 500 and 1000). In settings with weak overlap, the other methods overtake BART as the sample size increases to N = 5000 (Panels (D), (E) and (F)). Although they perform similarly, there is again evidence that the NDR-learner achieves the lowest RMSE of the policy advantage in larger samples.

[14] This is calculated as the advantage of a policy which treats everyone with a *true* negative CATE.

FIGURE 8. RMSE of ATEs: Rare Outcome Prevalence



6.4. **Performance: Plug-In Policy.** We now turn our attention to performance under the policy which simply treats everyone who has an estimated $\hat{\tau} < 0$. Tables 7 and 8 compare the NRMSE of this policy to the depth-two tree-based policy in the previous section. For normal outcome prevalence in settings with the lowest degree of heterogeneity (Setting 1), the tree-based policy achieves lower NRMSE (the second column from the left) than the plug-in policy (the first column). However, in Setting 1 with rare outcome prevalence, the plug-in policy performs better than the tree-based policy. With both normal and rare outcomes, the plug-in policy also performs better as heterogeneity increases - although the performance gap narrows significantly as sample size increases.

6.5. **Discussion of Findings.** We find that the various ML methods for obtaining double-robust scores used in policy learning perform very similarly in settings with normal outcome prevalence. However, our simulation results evidence that method performance diverges in settings with rare outcomes - BART seems to do very well in small sample sizes, and the NDR-learner does marginally better as N increases. As BART has not been adapted (to our knowledge) to the double-robust policy learning setting, we have shown evidence that this

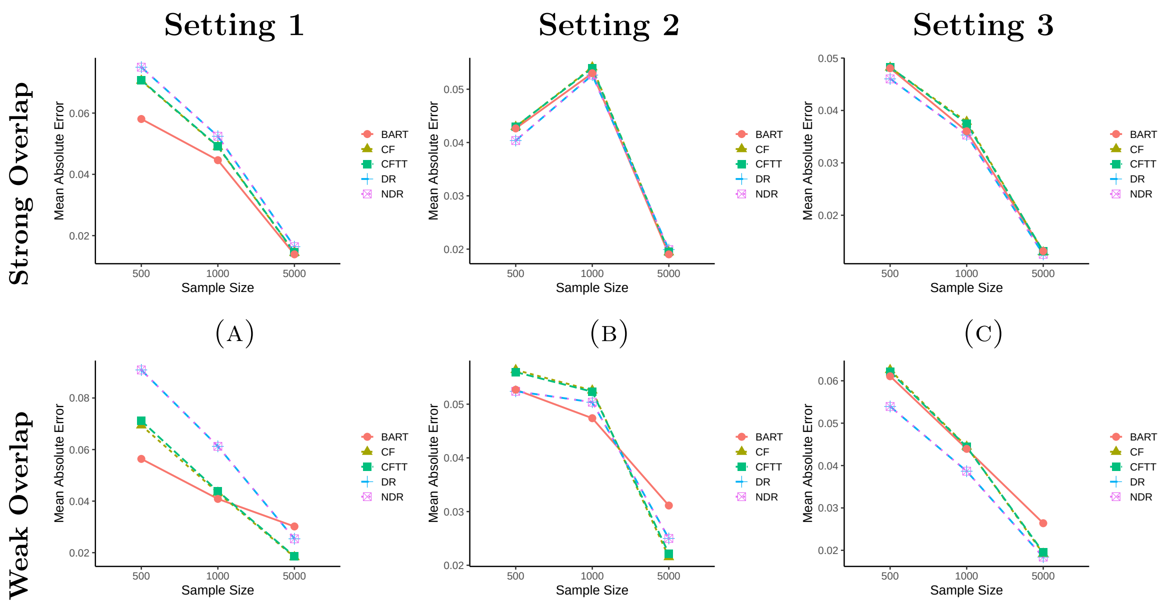
Table 4. RMSE of ATEs (SD)

Normal: Strong Overlap	SETTING 1			SETTING 2: SIM 7			SETTING 3: SIM 11		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.038	0.030	0.014	0.010	0.028	0.012	0.040	0.027	0.013
	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.001)	(0.000)
NDR	0.038	0.030	0.014	0.010	0.028	0.012	0.040	0.027	0.013
	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.001)	(0.000)
CF	0.039	0.029	0.013	0.012	0.030	0.012	0.042	0.029	0.013
	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.001)	(0.000)
CFTT	0.039	0.029	0.013	0.012	0.030	0.012	0.042	0.028	0.013
	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.001)	(0.000)
BART	0.043	0.033	0.013	0.012	0.034	0.013	0.046	0.030	0.014
	(0.003)	(0.001)	(0.000)	(0.000)	(0.002)	(0.000)	(0.004)	(0.001)	(0.000)
Weak Overlap	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.053	0.042	0.026	0.047	0.034	0.022	0.047	0.038	0.020
	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
NDR	0.053	0.042	0.026	0.047	0.034	0.022	0.047	0.038	0.020
	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
CF	0.054	0.044	0.023	0.056	0.040	0.021	0.057	0.043	0.021
	(0.004)	(0.003)	(0.001)	(0.005)	(0.003)	(0.001)	(0.005)	(0.003)	(0.001)
CFTT	0.053	0.043	0.023	0.055	0.040	0.021	0.057	0.043	0.021
	(0.004)	(0.003)	(0.001)	(0.005)	(0.003)	(0.001)	(0.005)	(0.003)	(0.001)
BART	0.071	0.057	0.040	0.062	0.046	0.036	0.061	0.050	0.035
	(0.007)	(0.005)	(0.002)	(0.006)	(0.004)	(0.002)	(0.005)	(0.004)	(0.002)
Rare: Strong Overlap	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.011	0.007	0.004	0.009	0.006	0.003	0.008	0.006	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NDR	0.011	0.007	0.004	0.009	0.006	0.003	0.008	0.006	0.003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CF	0.011	0.007	0.004	0.010	0.007	0.003	0.009	0.006	0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CFTT	0.012	0.007	0.004	0.009	0.007	0.003	0.009	0.006	0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BART	0.011	0.008	0.004	0.010	0.007	0.004	0.010	0.007	0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Weak Overlap	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.015	0.009	0.005	0.010	0.007	0.004	0.010	0.007	0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NDR	0.015	0.009	0.005	0.010	0.007	0.004	0.010	0.007	0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CF	0.010	0.010	0.008	0.012	0.009	0.005	0.013	0.008	0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CFTT	0.012	0.010	0.007	0.012	0.009	0.005	0.012	0.009	0.005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BART	0.021	0.013	0.010	0.012	0.010	0.007	0.012	0.009	0.008
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

[a] This table reports RMSE of estimated ATEs. The top two panels depict normal outcome prevalence, and the bottom two panels depict rare outcome prevalence.

ML method should be considered among possible candidates for estimating DR-scores used in optimal policy learning, especially for certain DGPs.

FIGURE 9. MAE of ATEs: Normal Outcome Prevalence



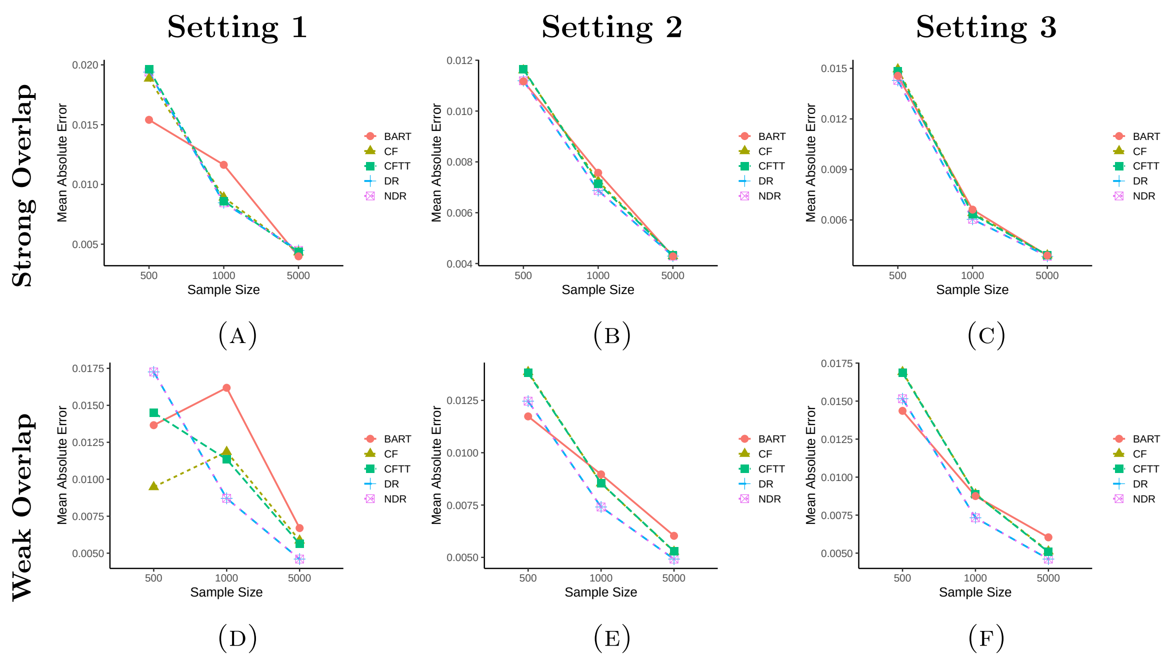
(A) (B) (C)

(D) (E) (F)

7. Robustness Checks

7.1. **Constant Propensity Score.** To examine whether the high levels of confounding in treatment assignment are influencing treatment effect and policy learning, we conduct a version of the simulations in which the true propensity score is a constant 0.30 across all observations. Table 9 depicts results for RMSE of the CATEs and ATE, and Table 10 presents NRMSE of Policy Advantage A_i. As in the main results, BART significantly outperforms the other methods in the case of rare outcomes. In settings with larger degrees of heterogeneity, BART retains its advantage in larger sample sizes, as well. We also note that the RMSE of the ATE for BART is no longer the worst performing. The continued good performance of BART for the policy advantage is likely due to BART's power in modeling response surfaces - the functional form for the probability of treatment perhaps matters less for this particular algorithm. In general, the NRMSEs for Settings 2 and 3 across all ML methods are slightly lower in magnitude than for the settings with weak overlap (confounding) and rare outcomes.

Figure 10. MAE of ATEs: Rare Outcome Prevalence



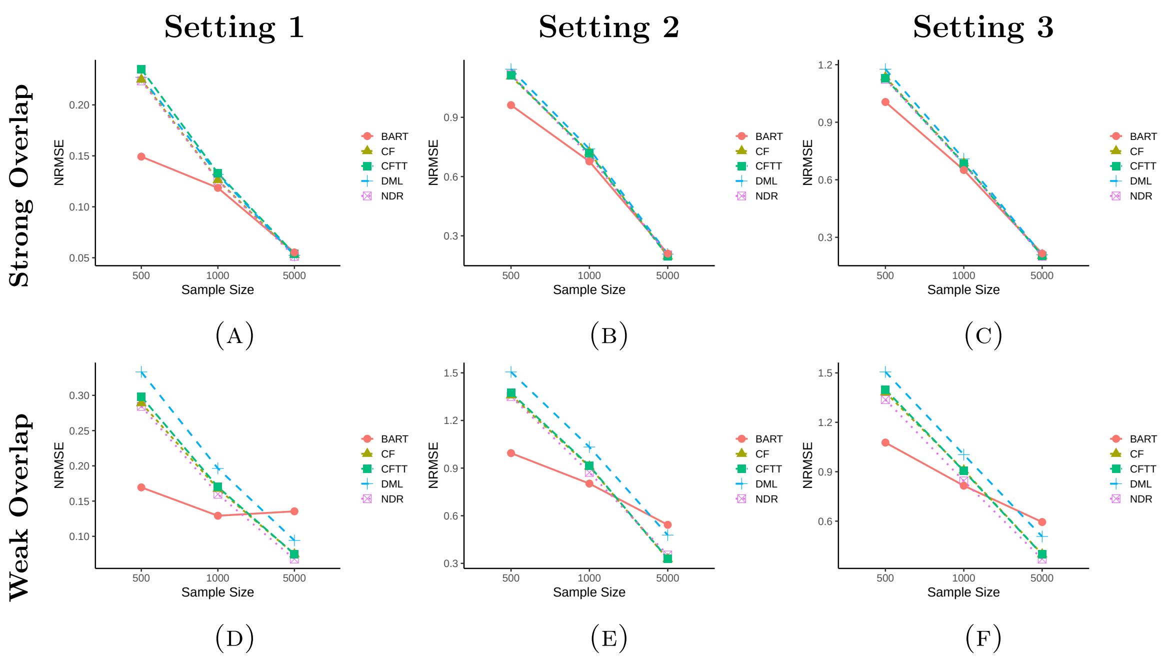
8. Case Study: Indonesia National Health Insurance Programme

We now return to our case study examining the impact of the Indonesian National Health Insurance Programme on infant mortality. This setting allows us to compare estimates across methods for a rare outcome, as mortality rates during the sample period were 2.6%.

8.1. **Results.** Table 11 shows results for the case study on subsidized health insurance and infant mortality. We report estimated ATE and the policy advantage (Equation 6) of a depth-two decision tree. Note that the DML policy seems to obtain the greatest reduction in infant mortality. However, results of our simulations indicate that the NDR-learner policy advantage is likely the more accurate estimate, especially given the ratio of covariates to observations in the data.

We therefore turn to the learned policy from our NDR-learner score estimates in order to examine characteristics of those who would be insured under the learned policy. Table 12 compares mean characteristics of those who were actually insured in the sample, those

Figure 11. NRMSE of Policy Advantage: Normal Outcome Prevalence



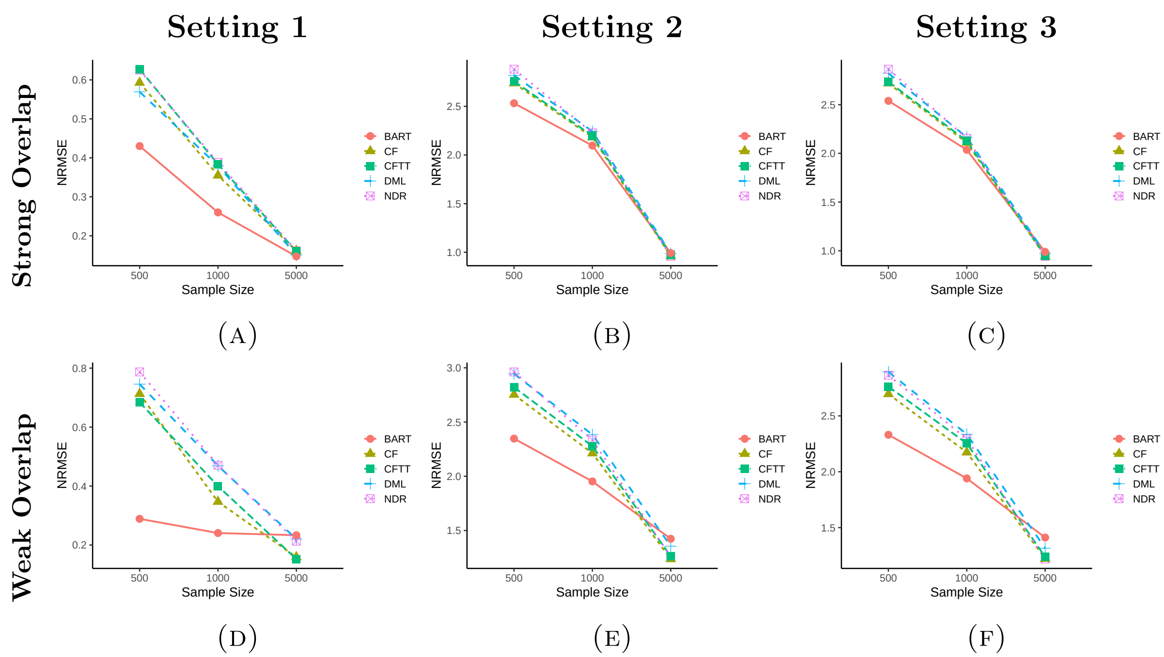
who would be newly insured under the depth-two NDR-learner policy, and the full sample.[15] We can see that the learned policy treats fewer individuals with the Poor Card (recall that in our heterogeneity analysis of BLP of GATEs, this subgroup had an undesirable positive signed treatment effect), and treats a larger portion of individuals ages 23-27, the subgroup exhibiting a significant portion of the negative treatment effect in our earlier analysis. This result demonstrates the potential for the learned policy in a simple tree-based class to point to drivers of treatment effect heterogeneity and account for them in the treatment assignment rule.

9. Conclusions

In this paper, we ask whether the choice of machine learning method for obtaining double-robust scores used to obtain particular types of treatment allocation rules matters in settings with rare outcomes. Indeed, although the ML methods we consider perform similarly in settings with normal outcome prevalence, we find significant differences in their performance

[15] Recall that, although the algorithm is trained on the full set of covariates, we restrict the policy class to only use the set of covariates explored in the GATE analysis.

FIGURE 12. NRMSE of Policy Advantage: Rare Outcome Prevalence



when rare outcomes are being targeted for intervention. Specifically, we find that using Bayesian Additive Regression Trees (BART) to obtain double-robust scores used in policy learning may improve performance in small samples. We also find that the NDR-learner of Knaus (2020) shows promise in this setting.

Our results also indicate that the adapted Causal Forest algorithm using testing and training (cross-fitting) samples may yield estimates which differ slightly from the usual Honest Causal Forest method. We have also shown that, in the case of rare outcomes, a simple plug-in policy may perform better in terms of Policy Advantage than a tree-based policy class. We leave discussion of equity and interpretability concerns regarding various policy classes to future work.

We note, however, that our simulation design only has 10 covariates: as the size of the sample increases, the ratio of covariates to observations becomes very small, which may not be similar to the structure of observational datasets. For example, our IFLS case study data contains 64 covariates and 10,622 observations, resulting in a ratio of 0.006 (x to n). The simulation setting with 1,000 observations has a ratio of 0.01, and for 5,000 observations a ratio of 0.002, implying that the larger simulation sample may be too low-dimensional to

Table 5. Weak Overlap: NRMSE of Policy Advantage (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DML	0.333	0.196	0.094	1.506	1.033	0.478	1.506	1.004	0.507
	(0.007)	(0.001)	(0.001)	(0.012)	(0.007)	(0.002)	(0.011)	(0.005)	(0.002)
NDR	0.284	0.160	**0.068**	1.351	0.873	0.353	1.337	0.844	**0.370**
	(0.006)	(0.001)	(0.001)	(0.011)	(0.005)	(0.001)	(0.010)	(0.005)	(0.001)
CF	0.289	0.168	0.075	1.359	0.910	**0.327**	1.382	0.909	0.399
	(0.006)	(0.001)	(0.001)	(0.011)	(0.006)	(0.001)	(0.011)	(0.006)	(0.002)
CFTT	0.298	0.170	0.075	1.374	0.915	0.329	1.399	0.904	0.397
	(0.007)	(0.001)	(0.001)	(0.010)	(0.006)	(0.001)	(0.011)	(0.005)	(0.002)
BART	**0.169**	**0.129**	0.135	**0.994**	**0.803**	0.543	**1.077**	**0.815**	0.594
	(0.003)	(0.002)	(0.002)	(0.009)	(0.007)	(0.003)	(0.010)	(0.006)	(0.004)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DML	0.746	0.469	0.220	2.945	2.384	1.355	2.894	2.336	1.315
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
NDR	0.787	0.471	0.213	2.962	2.334	1.258	2.864	2.294	**1.218**
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CF	0.713	0.347	0.159	2.752	2.212	**1.237**	2.696	2.174	1.221
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CFTT	0.685	0.400	**0.152**	2.821	2.278	1.262	2.762	2.254	1.237
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
BART	**0.289**	**0.241**	0.233	**2.347**	**1.953**	1.424	**2.331**	**1.940**	1.411
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)

[a] This table reports Normalized RMSE of estimated depth-two optimal policy for settings with poor overlap. RMSE is normalized by the mean advantage of the oracle policy. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

extrapolate to our setting. We caution against interpreting our findings for real-world data applications without considering the dimensions of the observational dataset in question.

Table 6. Strong Overlap: NRMSE of Policy Advantage (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DML	0.227	0.132	0.052	2.945	0.740	0.208	1.176	0.709	0.209
	(0.004)	(0.002)	(0.000)	(0.001)	(0.004)	(0.001)	(0.008)	(0.004)	(0.001)
NDR	0.224	0.125	0.051	2.962	0.705	0.203	1.124	0.681	0.201
	(0.004)	(0.002)	(0.000)	(0.001)	(0.004)	(0.001)	(0.008)	(0.004)	(0.001)
CF	0.225	0.126	0.053	2.752	0.722	0.196	1.135	0.687	0.203
	(0.004)	(0.002)	(0.000)	(0.001)	(0.004)	(0.001)	(0.008)	(0.004)	(0.001)
CFTT	0.235	0.133	0.054	2.821	0.718	0.197	1.129	0.685	0.204
	(0.004)	(0.002)	(0.000)	(0.001)	(0.004)	(0.001)	(0.008)	(0.004)	(0.001)
BART	0.149	0.119	0.055	2.347	0.677	0.210	1.005	0.651	0.216
	(0.003)	(0.001)	(0.000)	(0.001)	(0.004)	(0.001)	(0.007)	(0.004)	(0.001)

Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000			
DML	0.569	0.380	0.151	2.819	2.243	0.988	2.822	2.165	0.973
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
NDR	0.625	0.388	0.161	2.882	2.228	0.959	2.864	2.155	0.943
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CF	0.593	0.355	0.162	2.738	2.183	0.960	2.719	2.109	0.943
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CFTT	0.627	0.384	0.160	2.758	2.198	0.973	2.736	2.127	0.950
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
BART	0.430	0.260	0.147	2.532	2.097	0.992	2.539	2.036	0.988
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)

[a] This table reports Normalized RMSE of estimated depth-two optimal policy for settings with moderately strong overlap. RMSE is normalized by the mean advantage of the oracle policy. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

FIGURE 13. Estimated vs True Policy Advantage, Trees: Normal Outcome Prevalence

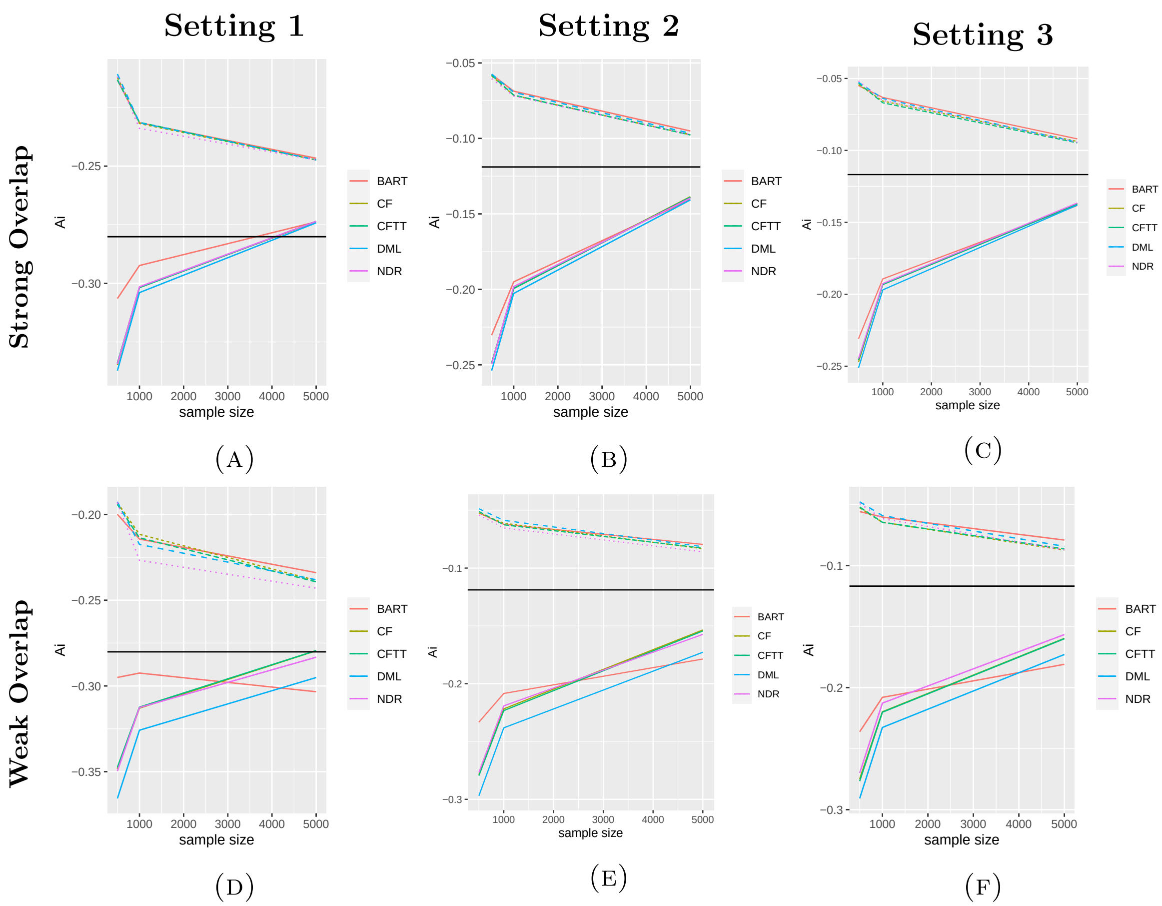


FIGURE 14. Estimated vs True Policy Advantage, Trees: Rare Outcome Prevalence

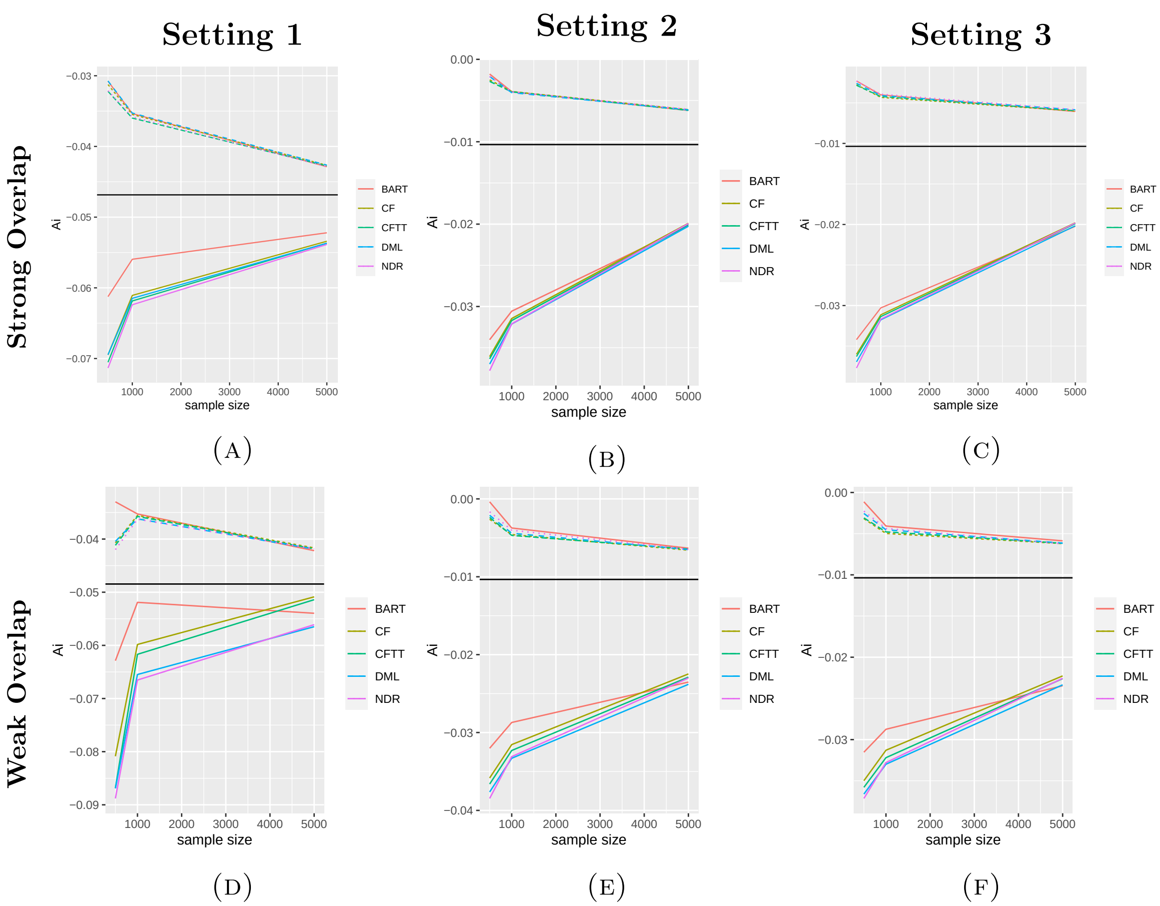


Table 7. Weak Overlap: Raw Estimated Policy Advantage, Trees (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
DML	−0.366	−0.326	−0.295	−0.297	−0.238	−0.173	−0.291	−0.233	−0.173
	(0.057)	(0.044)	(0.027)	(0.059)	(0.046)	(0.028)	(0.059)	(0.046)	(0.028)
NDR	−0.350	−0.313	−0.283	−0.277	−0.219	−0.157	−0.270	−0.213	−0.157
	(0.052)	(0.039)	(0.021)	(0.054)	(0.040)	(0.022)	(0.054)	(0.040)	(0.022)
CF	−0.348	−0.313	−0.279	−0.279	−0.222	−0.154	−0.275	−0.220	−0.160
	(0.053)	(0.040)	(0.022)	(0.054)	(0.041)	(0.022)	(0.054)	(0.041)	(0.022)
CFTT	−0.347	−0.312	−0.279	−0.280	−0.223	−0.154	−0.277	−0.220	−0.160
	(0.052)	(0.040)	(0.022)	(0.054)	(0.041)	(0.022)	(0.054)	(0.041)	(0.022)
BART	−0.295	−0.292	−0.303	−0.233	−0.208	−0.179	−0.236	−0.208	−0.181
	(0.046)	(0.040)	(0.035)	(0.048)	(0.042)	(0.034)	(0.047)	(0.041)	(0.034)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.051	−0.047	−0.047	−0.117	−0.117	−0.047	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DML	−0.087	−0.065	−0.056	−0.038	−0.033	−0.024	−0.037	−0.033	−0.023
	(0.016)	(0.011)	(0.006)	(0.012)	(0.009)	(0.005)	(0.012)	(0.009)	(0.005)
NDR	−0.089	−0.067	−0.056	−0.038	−0.033	−0.023	−0.037	−0.033	−0.023
	(0.015)	(0.011)	(0.005)	(0.012)	(0.009)	(0.005)	(0.012)	(0.009)	(0.005)
CF	−0.081	−0.060	−0.051	−0.036	−0.032	−0.022	−0.035	−0.031	−0.022
	(0.015)	(0.011)	(0.006)	(0.011)	(0.009)	(0.005)	(0.011)	(0.009)	(0.005)
CFTT	−0.087	−0.062	−0.051	−0.037	−0.032	−0.023	−0.036	−0.032	−0.023
	(0.015)	(0.011)	(0.006)	(0.011)	(0.009)	(0.005)	(0.011)	(0.009)	(0.005)
BART	−0.063	−0.052	−0.054	−0.032	−0.029	−0.024	−0.032	−0.029	−0.023
	(0.012)	(0.010)	(0.007)	(0.010)	(0.008)	(0.006)	(0.010)	(0.008)	(0.006)

[a] This table reports raw Ai of estimated depth-two optimal policy for settings with poor overlap. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

Table 8. Strong Overlap: Raw Estimated Policy Advantage, Trees (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
DML	−0.337	−0.304	−0.274	−0.254	−0.203	−0.141	−0.251	−0.197	−0.138
	(0.045)	(0.033)	(0.015)	(0.048)	(0.034)	(0.016)	(0.048)	(0.034)	(0.016)
NDR	−0.334	−0.301	−0.274	−0.250	−0.198	−0.140	−0.245	−0.193	−0.137
	(0.044)	(0.032)	(0.015)	(0.046)	(0.034)	(0.016)	(0.047)	(0.034)	(0.016)
CF	−0.335	−0.302	−0.274	−0.249	−0.199	−0.139	−0.247	−0.193	−0.137
	(0.044)	(0.032)	(0.015)	(0.047)	(0.033)	(0.016)	(0.047)	(0.033)	(0.016)
CFTT	−0.334	−0.302	−0.274	−0.249	−0.199	−0.139	−0.246	−0.193	−0.137
	(0.044)	(0.032)	(0.015)	(0.046)	(0.033)	(0.016)	(0.046)	(0.033)	(0.016)
BART	−0.307	−0.292	−0.274	−0.230	−0.195	−0.140	−0.231	−0.189	−0.138
	(0.040)	(0.031)	(0.016)	(0.043)	(0.033)	(0.017)	(0.043)	(0.033)	(0.017)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000			
OR	−0.046	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DML	−0.069	−0.061	−0.054	−0.037	−0.032	−0.020	−0.037	−0.032	−0.020
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
NDR	−0.071	−0.062	−0.054	−0.038	−0.032	−0.020	−0.038	−0.032	−0.020
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
CF	−0.069	−0.061	−0.053	−0.036	−0.031	−0.020	−0.036	−0.031	−0.020
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
CFTT	−0.071	−0.062	−0.054	−0.036	−0.032	−0.020	−0.036	−0.031	−0.020
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
BART	−0.061	−0.056	−0.052	−0.034	−0.031	−0.020	−0.034	−0.030	−0.020
	(0.013)	(0.010)	(0.005)	(0.010)	(0.008)	(0.004)	(0.010)	(0.008)	(0.004)

[a] This table reports Normalized RMSE of estimated depth-two optimal policy for settings with moderately strong overlap. RMSE is normalized by the mean advantage of the oracle policy. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

Table 9. Weak Overlap: True Policy Advantage, Trees (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
DML	−0.193	−0.217	−0.238	−0.049	−0.059	−0.082	−0.048	−0.059	−0.084
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.002)
NDR	−0.193	−0.227	−0.243	−0.054	−0.065	−0.086	−0.049	−0.062	−0.087
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.002)
CF	−0.193	−0.212	−0.238	−0.052	−0.061	−0.083	−0.053	−0.065	−0.087
	(0.011)	(0.008)	(0.003)	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.002)
CFTT	−0.194	−0.214	−0.239	−0.051	−0.063	−0.083	−0.052	−0.065	−0.086
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.002)
BART	−0.200	−0.214	−0.234	−0.052	−0.062	−0.080	−0.056	−0.060	−0.079
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.002)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.051	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DML	−0.040	−0.036	−0.042	−0.002	−0.004	−0.007	−0.003	−0.005	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
NDR	−0.042	−0.036	−0.042	−0.002	−0.004	−0.006	−0.002	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CF	−0.041	−0.036	−0.042	−0.003	−0.005	−0.007	−0.003	−0.005	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CFTT	−0.041	−0.036	−0.042	−0.002	−0.005	−0.006	−0.003	−0.005	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
BART	−0.033	−0.035	−0.042	−0.000	−0.004	−0.006	−0.001	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)

[a] This table reports true Ai of estimated depth-two optimal policy for settings with poor overlap. Advantage is calculated using the learned policies with the true CATE. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

Table 10. Strong Overlap: True Policy Advantage, Trees (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
DML	−0.211	−0.232	−0.247	−0.057	−0.069	−0.097	−0.053	−0.064	−0.094
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.001)	(0.006)	(0.004)	(0.001)
NDR	−0.212	−0.234	−0.247	−0.060	−0.072	−0.097	−0.052	−0.065	−0.094
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.001)	(0.006)	(0.004)	(0.001)
CF	−0.212	−0.232	−0.247	−0.059	−0.071	−0.098	−0.054	−0.066	−0.094
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.001)	(0.006)	(0.004)	(0.001)
CFTT	−0.213	−0.231	−0.247	−0.058	−0.071	−0.098	−0.053	−0.067	−0.095
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.001)	(0.006)	(0.004)	(0.001)
BART	−0.213	−0.232	−0.247	−0.058	−0.069	−0.095	−0.055	−0.063	−0.092
	(0.011)	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)	(0.006)	(0.004)	(0.001)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000			
OR	−0.046	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DML	−0.031	−0.035	−0.043	−0.002	−0.004	−0.006	−0.003	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
NDR	−0.032	−0.036	−0.043	−0.002	−0.004	−0.006	−0.002	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CF	−0.031	−0.035	−0.043	−0.002	−0.004	−0.006	−0.003	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CFTT	−0.032	−0.036	−0.043	−0.003	−0.004	−0.006	−0.003	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
BART	−0.031	−0.035	−0.043	−0.002	−0.004	−0.006	−0.002	−0.004	−0.006
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)

[a] This table reports the true advantages of estimated depth-two optimal policy for settings with moderately strong overlap. Advantage is calculated using the learned policies with the true CATE. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

Table 11. Comparison to Plug-In Policy: Normal Prevalence

SETTING 1	Strong				Weak			
Policy:	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree
N = 500	NRMSE	NRMSE	%	%	NRMSE	NRMSE	%	%
OR	0.000	0.000	0.442	0.442	0.000	0.000	0.442	0.442
NDR	0.258	0.224	0.497	0.476	0.376	0.284	0.533	0.495
CF	0.246	0.225	0.453	0.475	0.378	0.289	0.394	0.459
CFTT	0.243	0.235	0.456	0.474	0.364	0.298	0.391	0.461
BART	0.172	0.149	0.437	0.463	0.293	0.169	0.371	0.420
N = 1000								
OR	0.000	0.000	0.446	0.446	0.000	0.000	0.446	0.446
NDR	0.199	0.125	0.475	0.465	0.227	0.160	0.490	0.479
CF	0.174	0.126	0.443	0.466	0.255	0.168	0.370	0.443
CFTT	0.176	0.133	0.444	0.466	0.260	0.170	0.369	0.444
BART	0.134	0.119	0.443	0.454	0.203	0.129	0.390	0.412
N = 5000								
OR	0.000	0.000	0.447	0.447	0.000	0.000	0.447	0.447
NDR	0.077	0.051	0.455	0.458	0.121	0.068	0.469	0.470
CF	0.079	0.053	0.445	0.455	0.151	0.075	0.383	0.434
CFTT	0.077	0.054	0.445	0.454	0.148	0.075	0.383	0.435
BART	0.061	0.055	0.451	0.453	0.132	0.135	0.421	0.426
SETTING 2	Strong				Weak			
Policy:	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree
N = 500								
OR	0.000	0.000	0.866	0.866	0.000	0.000	0.866	0.866
NDR	0.472	1.118	0.941	0.729	0.721	1.351	0.899	0.715
CF	0.367	1.107	0.981	0.727	0.838	1.359	0.940	0.719
CFTT	0.467	1.114	0.962	0.725	0.676	1.374	0.909	0.712
BART	0.325	0.962	0.954	0.726	0.452	0.994	0.903	0.713
N = 1000								
OR	0.000	0.000	0.864	0.864	0.000	0.000	0.864	0.864
NDR	0.363	0.705	0.947	0.777	0.456	0.873	0.926	0.756
CF	0.358	0.722	0.984	0.778	0.410	0.910	0.963	0.746
CFTT	0.350	0.718	0.963	0.775	0.425	0.915	0.926	0.747
BART	0.303	0.677	0.949	0.772	0.360	0.803	0.920	0.749
N = 5000								
OR	0.000	0.000	0.866	0.866	0.000	0.000	0.866	0.866
NDR	0.148	0.203	0.975	0.877	0.246	0.353	0.944	0.853
CF	0.132	0.196	0.988	0.879	0.187	0.327	0.973	0.844
CFTT	0.141	0.197	0.975	0.880	0.225	0.329	0.950	0.836
BART	0.133	0.210	0.929	0.872	0.280	0.543	0.909	0.812
SETTING 3	Strong				Weak			
Policy:	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree
N = 500								
OR	0.000	0.000	0.901	0.901	0.000	0.000	0.901	0.901
NDR	0.597	1.124	0.905	0.697	0.751	1.337	0.894	0.700
CF	0.794	1.135	0.945	0.713	0.624	1.382	0.944	0.717
CFTT	0.577	1.129	0.932	0.711	0.697	1.399	0.913	0.717
BART	0.354	1.005	0.927	0.721	0.494	1.077	0.921	0.728
N = 1000								
OR	0.000	0.000	0.900	0.900	0.000	0.000	0.900	0.900
NDR	0.421	0.681	0.947	0.763	0.507	0.844	0.901	0.750
CF	0.333	0.687	0.990	0.767	0.392	0.909	0.973	0.770
CFTT	0.387	0.685	0.969	0.771	0.454	0.904	0.937	0.774
BART	0.270	0.651	0.955	0.758	0.324	0.815	0.931	0.746
N = 5000								
OR	0.000	0.000	0.900	0.900	0.000	0.000	0.900	0.900
NDR	0.150	0.201	0.981	0.906	0.243	0.370	0.949	0.879
CF	0.134	0.203	0.998	0.907	0.190	0.399	0.995	0.882
CFTT	0.135	0.204	0.992	0.911	0.198	0.397	0.982	0.879
BART	0.128	0.216	0.952	0.897	0.280	0.594	0.951	0.841

[a] This table reports the calculated advantage of a naive policy that treats all observations with a negative CATE, the percent of observations treated under the plug-in policy, and advantage and percent treated under the depth-two policy tree using all covariates as decision criteria. The outcome prevalence is normal. The left panel is settings with strong overlap, and the right panel is weak overlap.

FIGURE 15. Estimated vs True Policy Advantage, Plug-in: Normal Outcome Prevalence

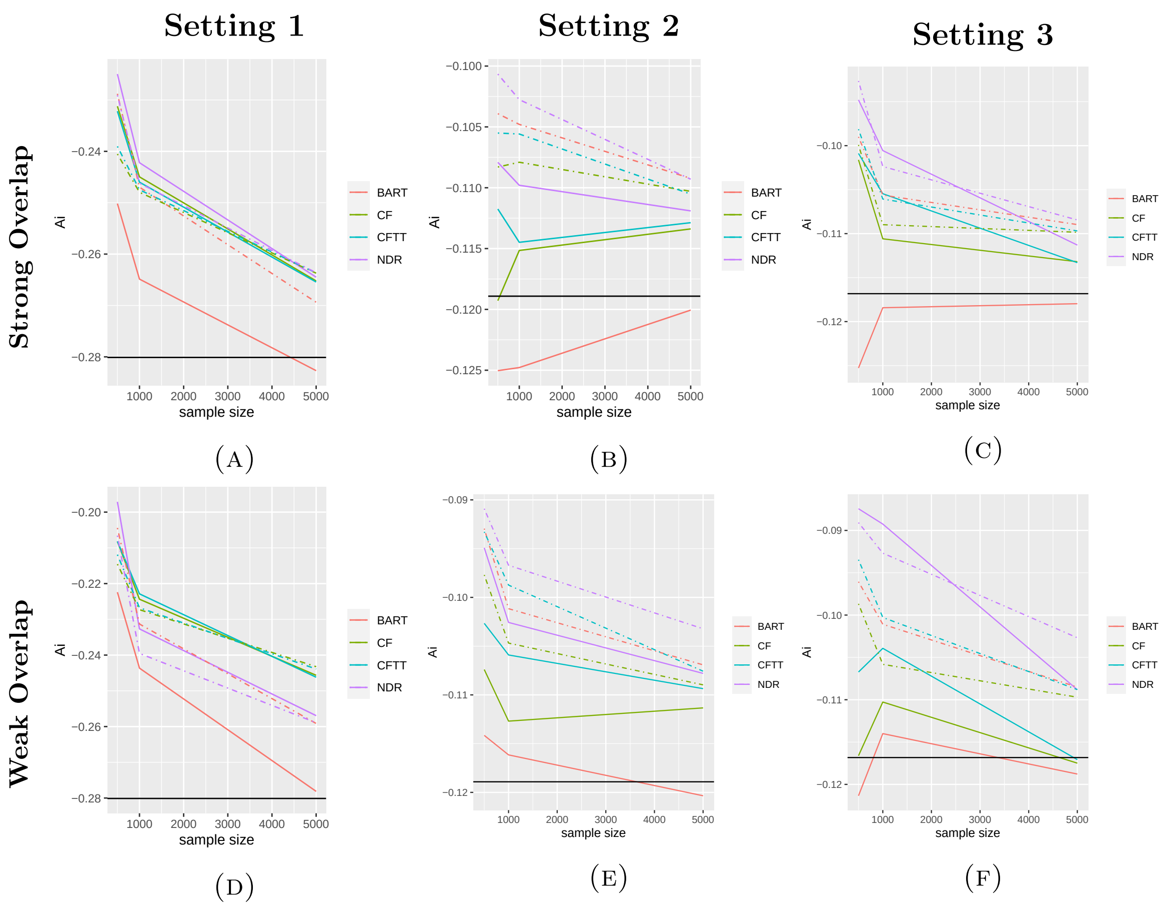


FIGURE 16. Estimated vs True Policy Advantage, Trees: Rare Outcome Prevalence

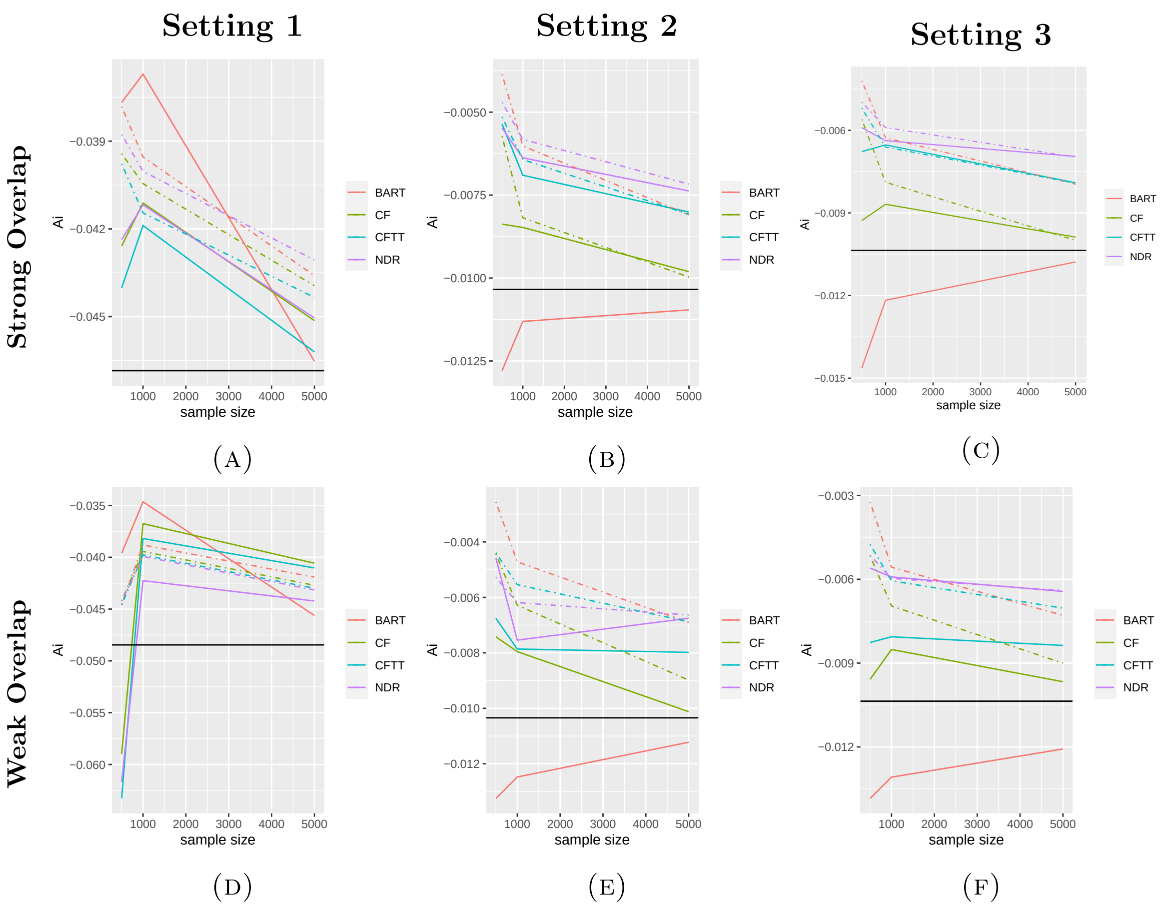


TABLE 12. Comparison to Plug-in Policy: Rare Prevalence

SETTING 1	Strong				Weak			
Policy:	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree
N = 500	NRMSE	NRMSE	%	%	NRMSE	NRMSE	%	%
OR	0.000	0.000	0.481	0.481	0.000	0.000	0.504	0.504
NDR	0.349	0.625	0.917	0.650	0.336	0.787	0.952	0.698
CF	0.375	0.593	0.886	0.700	0.248	0.713	0.959	0.663
CFTT	0.346	0.627	0.856	0.674	0.326	0.685	0.878	0.707
BART	0.346	0.430	0.928	0.627	0.289	0.289	0.922	0.665
N = 1000								
OR	0.000	0.000	0.484	0.484	0.000	0.000	0.484	0.484
NDR	0.247	0.388	0.892	0.688	0.290	0.471	0.854	0.721
CF	0.232	0.355	0.877	0.685	0.352	0.347	0.826	0.691
CFTT	0.241	0.384	0.790	0.699	0.347	0.400	0.768	0.687
BART	0.292	0.260	0.938	0.643	0.340	0.241	0.856	0.643
N = 5000								
OR	0.000	0.000	0.485	0.485	0.000	0.000	0.485	0.485
NDR	0.124	0.161	0.736	0.608	0.144	0.213	0.704	0.662
CF	0.112	0.162	0.693	0.614	0.197	0.159	0.637	0.629
CFTT	0.107	0.160	0.653	0.611	0.186	0.152	0.604	0.622
BART	0.117	0.147	0.670	0.600	0.182	0.233	0.555	0.579
SETTING 2	Strong				Weak			
Policy:	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree
N = 500								
OR	0.000	0.000	0.966	0.966	0.000	0.000	0.966	0.966
NDR	1.437	2.882	0.721	0.608	1.883	2.962	0.720	0.609
CF	1.516	2.738	0.776	0.629	1.998	2.752	0.713	0.626
CFTT	1.540	2.758	0.741	0.638	1.863	2.821	0.701	0.612
BART	0.984	2.532	0.684	0.591	1.000	2.347	0.636	0.538
N = 1000								
OR	0.000	0.000	0.967	0.967	0.000	0.000	0.967	0.967
NDR	1.116	2.228	0.771	0.697	1.223	2.334	0.764	0.691
CF	1.265	2.183	0.897	0.695	1.601	2.212	0.806	0.706
CFTT	1.180	2.198	0.802	0.698	1.356	2.278	0.756	0.706
BART	0.854	2.097	0.786	0.689	0.882	1.953	0.733	0.681
N = 5000								
OR	0.000	0.000	0.967	0.967	0.000	0.000	0.967	0.967
NDR	0.586	0.959	0.827	0.778	0.777	1.258	0.783	0.779
CF	0.535	0.960	0.980	0.776	0.732	1.237	0.930	0.777
CFTT	0.574	0.973	0.870	0.781	0.678	1.262	0.807	0.768
BART	0.391	0.992	0.874	0.784	0.812	1.424	0.820	0.760
SETTING 3	Strong				Weak			
Policy:	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree	$\tau < 0$	Tree
N = 500								
OR	0.000	0.000	0.974	0.974	0.000	0.000	0.974	0.974
NDR	1.532	2.864	0.735	0.607	1.892	2.864	0.740	0.604
CF	1.564	2.719	0.770	0.630	1.826	2.696	0.748	0.638
CFTT	1.457	2.736	0.746	0.634	1.756	2.762	0.721	0.633
BART	0.946	2.539	0.695	0.608	1.049	2.331	0.650	0.549
N = 1000								
OR	0.000	0.000	0.975	0.975	0.000	0.000	0.975	0.975
NDR	1.129	2.155	0.774	0.685	1.241	2.294	0.776	0.696
CF	1.268	2.109	0.880	0.703	1.543	2.174	0.833	0.726
CFTT	1.231	2.127	0.807	0.698	1.321	2.254	0.780	0.716
BART	0.762	2.036	0.793	0.689	0.934	1.940	0.761	0.689
N = 5000								
OR	0.000	0.000	0.975	0.975	0.000	0.000	0.975	0.975
NDR	0.611	0.943	0.821	0.767	0.771	1.218	0.793	0.788
CF	0.419	0.943	0.982	0.775	0.764	1.221	0.932	0.786
CFTT	0.552	0.950	0.865	0.774	0.600	1.237	0.822	0.785
BART	0.395	0.988	0.865	0.780	0.742	1.411	0.836	0.772

[a] This table reports the calculated advantage of a naive policy that treats all observations with a negative CATE, the percent of observations treated under the plug-in policy, percent treated under the depth-two policy tree using all covariates as decision criteria. The outcome prevalence is rare. The left panels are settings with strong overlap, and the right panel is weak overlap.

Table 13. Strong Overlap: Estimated Raw Policy Advantage, Plug-in (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
NDR	−0.225	−0.242	−0.264	−0.108	−0.110	−0.112	−0.095	−0.101	−0.111
	(0.046)	(0.032)	(0.015)	(0.047)	(0.034)	(0.016)	(0.048)	(0.034)	(0.016)
CF	−0.231	−0.245	−0.265	−0.119	−0.115	−0.113	−0.102	−0.111	−0.113
	(0.046)	(0.032)	(0.015)	(0.048)	(0.034)	(0.016)	(0.048)	(0.034)	(0.016)
CFTT	−0.232	−0.246	−0.265	−0.112	−0.114	−0.113	−0.101	−0.105	−0.113
	(0.045)	(0.032)	(0.015)	(0.047)	(0.034)	(0.016)	(0.047)	(0.034)	(0.016)
BART	−0.250	−0.265	−0.283	−0.125	−0.125	−0.120	−0.125	−0.118	−0.118
	(0.041)	(0.031)	(0.016)	(0.044)	(0.033)	(0.017)	(0.044)	(0.033)	(0.017)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.046	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NDR	−0.042	−0.041	−0.045	−0.005	−0.006	−0.007	−0.006	−0.006	−0.007
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
CF	−0.043	−0.041	−0.045	−0.008	−0.008	−0.010	−0.009	−0.009	−0.010
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
CFTT	−0.044	−0.042	−0.046	−0.005	−0.007	−0.008	−0.007	−0.007	−0.008
	(0.015)	(0.010)	(0.005)	(0.011)	(0.008)	(0.004)	(0.011)	(0.008)	(0.004)
BART	−0.038	−0.037	−0.047	−0.013	−0.011	−0.011	−0.015	−0.012	−0.011
	(0.013)	(0.010)	(0.005)	(0.010)	(0.008)	(0.004)	(0.010)	(0.008)	(0.004)

[a] This table reports raw estiamted Ai of estimated plug-in optimal policy for settings with strong overlap. Advantage is calculated using the learned policies with the learned CATE. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

Table 14. Weak Overlap: Estimated Raw Policy Advantage, Plug-in (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
NDR	−0.197	−0.233	−0.257	−0.095	−0.103	−0.108	−0.087	−0.089	−0.109
	(0.053)	(0.039)	(0.021)	(0.055)	(0.041)	(0.022)	(0.055)	(0.041)	(0.022)
CF	−0.208	−0.224	−0.246	−0.107	−0.113	−0.111	−0.117	−0.110	−0.117
	(0.055)	(0.041)	(0.022)	(0.056)	(0.042)	(0.022)	(0.055)	(0.042)	(0.022)
CFTT	−0.208	−0.223	−0.246	−0.103	−0.106	−0.109	−0.107	−0.104	−0.117
	(0.054)	(0.041)	(0.022)	(0.055)	(0.042)	(0.022)	(0.055)	(0.041)	(0.022)
BART	−0.222	−0.244	−0.278	−0.114	−0.116	−0.120	−0.121	−0.114	−0.119
	(0.046)	(0.040)	(0.035)	(0.049)	(0.042)	(0.034)	(0.048)	(0.042)	(0.034)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.051	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NDR	−0.062	−0.042	−0.044	−0.005	−0.008	−0.007	−0.006	−0.006	−0.006
	(0.016)	(0.011)	(0.005)	(0.012)	(0.009)	(0.005)	(0.012)	(0.009)	(0.005)
CF	−0.059	−0.037	−0.041	−0.007	−0.008	−0.010	−0.010	−0.009	−0.010
	(0.015)	(0.011)	(0.006)	(0.011)	(0.009)	(0.005)	(0.011)	(0.009)	(0.005)
CFTT	−0.063	−0.038	−0.041	−0.007	−0.008	−0.008	−0.008	−0.008	−0.008
	(0.015)	(0.011)	(0.006)	(0.011)	(0.009)	(0.005)	(0.011)	(0.009)	(0.005)
BART	−0.040	−0.035	−0.046	−0.013	−0.012	−0.011	−0.014	−0.013	−0.012
	(0.013)	(0.010)	(0.007)	(0.010)	(0.008)	(0.006)	(0.010)	(0.008)	(0.006)

[a] This table reports raw estiamted Ai of estimated plug-in optimal policy for settings with poor overlap. Advantage is calculated using the learned policies with the learned CATE. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

TABLE 15. Strong Overlap: True Policy Advantage, Plug-in (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
NDR	−0.229	−0.246	−0.264	−0.101	−0.103	−0.109	−0.095	−0.101	−0.111
	(0.010)	(0.006)	(0.003)	(0.004)	(0.003)	(0.001)	(0.048)	(0.034)	(0.016)
CF	−0.241	−0.248	−0.264	−0.108	−0.108	−0.110	−0.102	−0.111	−0.113
	(0.009)	(0.006)	(0.003)	(0.004)	(0.003)	(0.001)	(0.048)	(0.034)	(0.016)
CFTT	−0.239	−0.247	−0.264	−0.106	−0.106	−0.111	−0.101	−0.105	−0.113
	(0.009)	(0.006)	(0.003)	(0.004)	(0.003)	(0.001)	(0.047)	(0.034)	(0.016)
BART	−0.229	−0.247	−0.269	−0.104	−0.105	−0.109	−0.125	−0.118	−0.118
	(0.010)	(0.006)	(0.002)	(0.004)	(0.003)	(0.001)	(0.044)	(0.033)	(0.017)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.046	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NDR	−0.042	−0.041	−0.045	−0.005	−0.006	−0.007	−0.005	−0.006	−0.007
	(0.015)	(0.010)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CF	−0.043	−0.041	−0.045	−0.006	−0.008	−0.010	−0.006	−0.008	−0.010
	(0.015)	(0.010)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CFTT	−0.044	−0.042	−0.046	−0.005	−0.006	−0.008	−0.005	−0.007	−0.008
	(0.015)	(0.010)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BART	−0.038	−0.037	−0.047	−0.004	−0.006	−0.008	−0.004	−0.006	−0.008
	(0.013)	(0.010)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

[a] This table reports true Ai of estimated plug-in optimal policy for settings with strong overlap. Advantage is calculated using the learned policies with the true CATE. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

Table 16. Weak Overlap: True Policy Advantage, Plug-in (SD)

Normal	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.280	−0.280	−0.281	−0.119	−0.118	−0.119	−0.117	−0.117	−0.117
	(0.007)	(0.005)	(0.002)	(0.004)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)
NDR	−0.207	−0.240	−0.259	−0.091	−0.097	−0.103	−0.089	−0.093	−0.103
	(0.011)	(0.007)	(0.003)	(0.005)	(0.003)	(0.001)	(0.005)	(0.003)	(0.001)
CF	−0.215	−0.227	−0.243	−0.098	−0.105	−0.109	−0.099	−0.106	−0.110
	(0.010)	(0.007)	(0.003)	(0.004)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)
CFTT	−0.212	−0.227	−0.244	−0.093	−0.099	−0.108	−0.093	−0.100	−0.109
	(0.011)	(0.007)	(0.003)	(0.005)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)
BART	−0.204	−0.231	−0.259	−0.093	−0.101	−0.107	−0.096	−0.101	−0.109
	(0.011)	(0.007)	(0.003)	(0.004)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)
Rare	SETTING 1			SETTING 2			SETTING 3		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
OR	−0.051	−0.047	−0.047	−0.010	−0.010	−0.010	−0.010	−0.010	−0.010
	(0.005)	(0.003)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NDR	−0.044	−0.040	−0.043	−0.005	−0.006	−0.007	−0.005	−0.006	−0.006
	(0.005)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CF	−0.044	−0.039	−0.043	−0.004	−0.006	−0.009	−0.005	−0.007	−0.009
	(0.005)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CFTT	−0.045	−0.040	−0.043	−0.004	−0.006	−0.007	−0.005	−0.006	−0.007
	(0.005)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BART	−0.045	−0.039	−0.042	−0.003	−0.005	−0.007	−0.003	−0.006	−0.007
	(0.005)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

[a] This table reports true Ai of estimated plug-in optimal policy for settings with poor overlap. Advantage is calculated using the learned policies with the true CATE. The top panel depicts normal outcome prevalence, and the bottom panel shows rare outcome prevalence.

TABLE 17. Constant Propensity Score

Rare Prevalence						
Setting 1	CATE: RMSE (SD)			ATE: NRMSE (SD)		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.011	0.008	0.003	0.011	0.007	0.003
	(0.008)	(0.006)	(0.002)	(0.000)	(0.000)	(0.000)
NDR	0.011	0.008	0.003	0.011	0.007	0.003
	(0.008)	(0.006)	(0.002)	(0.000)	(0.000)	(0.000)
CF	0.011	0.008	0.003	0.011	0.007	0.003
	(0.008)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
CFTT	0.011	0.008	0.003	0.011	0.007	0.003
	(0.008)	(0.005)	(0.003)	(0.000)	(0.000)	(0.000)
BART	0.013	0.010	0.004	0.011	0.008	0.003
	(0.010)	(0.007)	(0.003)	(0.000)	(0.000)	(0.000)
Setting 2	CATE: RMSE (SD)			ATE: NRMSE (SD)		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.010	0.006	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
NDR	0.010	0.006	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
CF	0.010	0.006	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
CFTT	0.010	0.006	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
BART	0.009	0.006	0.003	0.010	0.006	0.003
	(0.006)	(0.004)	(0.002)	(0.000)	(0.000)	(0.000)
Setting 3	CATE: RMSE (SD)			ATE: NRMSE (SD)		
	N = 500	N = 1000	N = 5000	N = 500	N = 1000	N = 5000
DR	0.010	0.006	0.003	0.009	0.006	0.003
	(0.008)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
NDR	0.010	0.006	0.003	0.009	0.006	0.003
	(0.008)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
CF	0.010	0.006	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
CFTT	0.010	0.006	0.003	0.009	0.006	0.003
	(0.007)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
BART	0.008	0.005	0.003	0.009	0.006	0.003
	(0.006)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)

[a] This table reports RMSE of estimated CATEs and ATEs for settings with rare outcome prevalence, using a constant propensity score $e(X_i) = 0.3$.

Table 18. Contant Propensity Score: NRMSE of Policy Advantage (SD)

	Setting 1			Setting 2			Setting 3		
	N = 500	N = 2000	N = 5000	N = 500	N = 2000	N = 5000	N = 500	N = 2000	N = 5000
DML	0.494	0.357	0.133	2.675	2.185	0.957	2.696	2.181	0.932
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
NDR	0.528	0.344	0.140	2.732	2.178	0.956	2.757	2.189	0.932
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CF	0.533	0.341	0.149	2.694	2.204	0.960	2.706	2.205	0.936
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
CFTT	0.557	0.366	0.143	2.727	2.207	0.962	2.734	2.211	0.935
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
BART	0.451	0.288	0.143	2.526	2.058	0.940	2.539	2.062	0.917
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)

[a] This table reports RMSE of estimated CATEs and NRMSE of estimated ATEs. The setting is rare outcomes and the probability of treatment is a constant 0.30 regardless of covariates.

Table 19. ATE: Insurance and Infant Mortality, IFLS

	Estimated ATE		**Policy Advantage**	
	ATE	(SE)	Ai	(SE)
Kreif et al (2021)	-0.005	(0.005)		
DML	-0.004	(0.005)	-0.023	(0.005)
NDR	-0.003	(0.006)	-0.021	(0.005)
CF	-0.005	(0.005)	-0.020	(0.005)
CFTT	-0.004	(0.005)	-0.019	(0.005)
BART	-0.003	(0.005)	-0.019	(0.005)

Table 20. Characteristics of Newly Insured Under Depth-2 Policy (NDR Learner)

	Previously Insured N = 1,511	Newly Insured N = 6,788	Full Sample N = 10,622
Rural	0.473	0.507	0.483
	(0.499)	(0.500)	(0.500)
Poor Card	0.201	0.006	0.107
	(0.401)	(0.076)	(0.309)
Disaster	0.277	0.238	0.239
	(0.448)	(0.426)	(0.427)
Cash	0.451	0.269	0.265
	(0.498)	(0.444)	(0.441)
Raskin Programme	0.716	0.505	0.530
	(0.451)	(0.500)	(0.499)
Literacy	0.950	0.950	0.955
	(0.217)	(0.218)	(0.206)
Midwife	0.831	0.822	0.821
	(0.375)	(0.382)	(0.384)
Secondary Education	0.275	0.257	0.258
	(0.447)	(0.437)	(0.438)
Senior Education	0.291	0.310	0.325
	(0.454)	(0.463)	(0.468)
Higher Education	0.056	0.090	0.090
	(0.229)	(0.287)	(0.287)
Wealth Quintile 2	0.252	0.251	0.219
	(0.434)	(0.434)	(0.414)
Wealth Quintile 3	0.214	0.273	0.222
	(0.411)	(0.445)	(0.416)
Wealth Quintile 4	0.160	0.031	0.197
	(0.367)	(0.174)	(0.398)
Wealth Quintile 5	0.077	0.215	0.154
	(0.267)	(0.411)	(0.361)
Age:23-27	0.194	0.252	0.243
	(0.395)	(0.434)	(0.429)
Age:27-31	0.214	0.212	0.215
	(0.410)	(0.409)	(0.411)
Age:Over 31	0.306	0.228	0.241
	(0.461)	(0.420)	(0.428)
Health	0.834	0.864	0.861
	(0.372)	(0.342)	(0.346)

References

Athey, S., J. Tibshirani, and S. Wager (2019). "Generalized random forests". In: *The Annals of Statistics* 47.2, pp. 1148–1178.

Athey, S. and S. Wager (2019). "Estimating treatment effects with causal forests: An application". In: *Observational Studies* 5.2, pp. 37–51.

– (2021). "Policy learning with observational data". In: *Econometrica* 89.1, pp. 133–161.

Bertsimas, D. and N. Kallus (2020). "From predictive to prescriptive analytics". In: *Management Science* 66.3, pp. 1025–1044.

Buja, A., T. Hastie, and R. Tibshirani (1989). "Linear smoothers and additive models". In: *The Annals of Statistics*, pp. 453–510.

Carnegie, N., V. Dorie, and J. L. Hill (2019). "Examining treatment effect heterogeneity using BART". In: *Observational Studies* 5.2, pp. 52–70.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, and W. Newey (2017). "Double/debiased/neyman machine learning of treatment effects". In: *American Economic Review* 107.5, pp. 261–65.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). *Double/debiased machine learning for treatment and structural parameters.*

Chernozhukov, V., M. Demirer, et al. (2018). *Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India.* Tech. rep. National Bureau of Economic Research.

Chipman, H. A., E. I. George, and R. E. McCulloch (2010). "BART: Bayesian additive regression trees". In: *The Annals of Applied Statistics* 4.1, pp. 266–298.

Franklin, J. M. et al. (2017). "Comparing the performance of propensity score methods in healthcare database studies with rare outcomes". In: *Statistics in medicine* 36.12, pp. 1946–1963.

Hahn, P. R., J. S. Murray, and C. M. Carvalho (2020). "Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion)". In: *Bayesian Analysis* 15.3, pp. 965–1056.

Hill, J., A. Linero, and J. Murray (2020). "Bayesian additive regression trees: a review and look forward". In: *Annual Review of Statistics and Its Application* 7, pp. 251–278.

Hill, J. L. (2011). "Bayesian nonparametric modeling for causal inference". In: *Journal of Computational and Graphical Statistics* 20.1, pp. 217–240.

Hirano, K. and J. R. Porter (2009). “Asymptotics for statistical treatment rules”. In: *Econometrica* 77.5, pp. 1683–1701.

Hu, L. and C. Gu (2021). “Estimation of causal effects of multiple treatments in healthcare database studies with rare outcomes”. In: *Health Services and Outcomes Research Methodology*, pp. 1–22.

Hu, L., C. Gu, et al. (2020). “Estimation of causal effects of multiple treatments in observational studies with a binary outcome”. In: *Statistical methods in medical research* 29.11, pp. 3218–3234.

Hu, L., J. Ji, and F. Li (2021). “Estimating heterogeneous survival treatment effect in observational data using machine learning”. In: *Statistics in Medicine*.

Kallus, N. and A. Zhou (2021). “Minimax-optimal policy learning under unobserved confounding”. In: *Management Science* 67.5, pp. 2870–2890.

Kennedy, E. H. (2020). “Optimal doubly robust estimation of heterogeneous causal effects”. In: *arXiv preprint arXiv:2004.14497*.

Kitagawa, T. and A. Tetenov (2018). “Who should be treated? empirical welfare maximization methods for treatment choice”. In: *Econometrica* 86.2, pp. 591–616.

Knaus, M. C. (2020). “Double machine learning based program evaluation under unconfoundedness”. In: *arXiv preprint arXiv:2003.03191*.

Kreif, N. et al. (2021). “Estimating heterogeneous policy impacts using causal machine learning: a case study of health insurance reform in Indonesia”. In: *Health Services and Outcomes Research Methodology*, pp. 1–36.

Künzel, S. R. et al. (2019). “Metalearners for estimating heterogeneous treatment effects using machine learning”. In: *Proceedings of the national academy of sciences* 116.10, pp. 4156–4165.

Luedtke, A. R. and M. J. van der Laan (2016). “Super-learning of an optimal dynamic treatment rule”. In: *The international journal of biostatistics* 12.1, pp. 305–332.

Luedtke, A. R. and M. J. Van Der Laan (2016). “Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy”. In: *Annals of statistics* 44.2, p. 713.

Manski, C. F. (2004). “Statistical treatment rules for heterogeneous populations”. In: *Econometrica* 72.4, pp. 1221–1246.

– (2009). *Identification for prediction and decision*. Harvard University Press.

Murphy, S. A. (2003). "Optimal dynamic treatment regimes". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65.2, pp. 331–355.

Qian, M. and S. A. Murphy (2011). "Performance guarantees for individualized treatment rules". In: *Annals of statistics* 39.2, p. 1180.

Robins, J. M. and A. Rotnitzky (1995). "Semiparametric efficiency in multivariate regression models with missing data". In: *Journal of the American Statistical Association* 90.429, pp. 122–129.

Rubin, D. B. (1974). "Estimating causal effects of treatments in randomized and nonrandomized studies." In: *Journal of educational Psychology* 66.5, p. 688.

Strauss, J., F. Witoelar, and B. Sikoki (2016). *The fifth wave of the Indonesia family life survey: overview and field report*. Vol. 1. Rand Santa Monica, CA, USA.

Tan, Y. V. and J. Roy (2019). "Bayesian additive regression trees and the General BART model". In: *Statistics in medicine* 38.25, pp. 5048–5069.

Tibshirani, J. et al. (2021). "Package 'grf'". In.

Van der Laan, M. J., S. Rose, et al. (2011). *Targeted learning: causal inference for observational and experimental data*. Vol. 4. Springer.

Wager, S. and S. Athey (2018). "Estimation and inference of heterogeneous treatment effects using random forests". In: *Journal of the American Statistical Association* 113.523, pp. 1228–1242.

Appendix A. Tables

Table 21. Data Properties, n = 10,000

PANEL A: STRONG OVERLAP

		Prevalence = normal			Prevalence = rare		
		All	Y = 0	Y = 1	All	Y = 0	Y = 1
Setting 1	Control (%)	63.26	34.93	28.33	63.58	60.18	3.4
	Treat	36.74	20.96	15.78	36.42	35.96	0.46
Setting 2	Control	63.08	28.45	34.63	64.48	63.1	1.38
	Treat	36.92	21.57	15.35	35.52	35.06	0.46
Setting 3	Control	63.52	29.37	34.15	64.01	62.54	1.47
	Treat	36.48	20.55	15.93	35.99	35.62	0.37

PANEL B: WEAK OVERLAP

		Prevalence = normal			Prevalence = rare		
		All	Y = 0	Y = 1	All	Y = 0	Y = 1
Setting 1	Control (%)	65.39	33	32.39	65.78	61.58	4.2
	Treat	34.61	19.79	14.82	34.22	33.84	0.38
Setting 2	Control	65.35	28.9	36.45	65.88	64.5	1.38
	Treat	34.65	19.68	14.97	34.12	33.69	0.43
Setting 3	Control	65.93	30.08	35.85	65.74	64.35	1.39
	Treat	34.07	19.51	14.56	34.26	33.92	0.34

[a] This table reports average treatment prevalence, and outcome prevalence by treatment group for each simulation scenario (for simulated datasets of 10,000 observations).