# Experimental Evaluation of Individualized Treatment Rules*

Kosuke Imai[†] Michael Lingzhi Li[‡]

First version: May 15, 2019
This version: May 5, 2021

**Abstract**

The increasing availability of individual-level data has led to numerous applications of individualized (or personalized) treatment rules (ITRs). Policy makers often wish to empirically evaluate ITRs and compare their relative performance before implementing them in a target population. We propose a new evaluation metric, the population average prescriptive effect (PAPE). The PAPE compares the performance of ITR with that of non-individualized treatment rule, which randomly treats the same proportion of units. Averaging the PAPE over a range of budget constraints yields our second evaluation metric, the area under the prescriptive effect curve (AUPEC). The AUPEC represents an overall performance measure for evaluation, like the area under the receiver and operating characteristic curve (AUROC) does for classification, and is a generalization of the QINI coefficient utilized in uplift modeling. We use Neyman's repeated sampling framework to estimate the PAPE and AUPEC and derive their exact finite-sample variances based on random sampling of units and random assignment of treatment. We extend our methodology to a common setting, in which the same experimental data is used to both estimate and evaluate ITRs. In this case, our variance calculation incorporates the additional uncertainty due to random splits of data used for cross-validation. The proposed evaluation metrics can be estimated without requiring modeling assumptions, asymptotic approximation, or resampling methods. As a result, it is applicable to any ITR including those based on complex machine learning algorithms. The open-source software package is available for implementing the proposed methodology.

**Key Words:** causal inference, heterogenous treatment effects, machine learning, precision medicine, uplift modeling

---

*We thank Naoki Egami, Colin Fogarty, Zhichao Jiang, Susan Murphy, Nicole Pashley, Stefan Wager, three anonymous reviewers, and especially the Associate Editor for many helpful comments. We thank Dr. Hsin-Hsiao Wang who inspired us to write this paper. Imai thanks the Alfred P. Sloan Foundation for partial support of this research. The proposed methodology is implemented through an open-source R package, evalITR, which is freely available for download at the Comprehensive R Archive Network (CRAN; `https://CRAN.R-project.org/package=evalITR`).

[†]Professor, Department of Government and Department of Statistics, Harvard University, Cambridge, MA 02138. Phone: 617–384–6778, Email: Imai@Harvard.Edu, URL: https://imai.fas.harvard.edu

[‡]Operation Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139. mlli@mit.edu

# 1 Introduction

In today's data-rich society, the individualized (or personalized) treatment rules (ITRs), which assign different treatments to individuals based on their observed characteristics, play an essential role. Examples include personalized medicine and micro-targeting in business and political campaigns (e.g., Hamburg and Collins, 2010; Imai and Strauss, 2011). In the causal inference literature, a number of researchers have developed methods to estimate optimal ITRs using a variety of machine learning algorithms (see e.g., Qian and Murphy, 2011; Zhang *et al.*, 2012; Fu *et al.*, 2016; Luedtke and van der Laan, 2016a,b; Zhou *et al.*, 2017; Athey and Wager, 2018; Kitagawa and Tetenov, 2018). In addition, applied researchers often use machine learning algorithms to estimate heterogeneous treatment effects and then construct ITRs based on the resulting estimates.

In this paper, we consider a common setting, in which a policy-maker wishes to experimentally evaluate the empirical performance of an ITR before implementing it in a target population. Such evaluation is also essential for comparing the efficacy of alternative ITRs. Specifically, we show how to use a randomized experiment for evaluating ITRs. We propose two new evaluation metrics. The first is the population average prescriptive effect (PAPE), which compares an ITR with a non-individualized treatment rule that randomly assigns the same proportion of units to the treatment condition. The PAPE represents the difference between the average outcome under the ITR and that under the random treatment rule. The key idea is that a well-performing ITR should outperform the random treatment rule, which does not utilize any individual-level information.

Averaging the PAPE over a range of budget constraints yields our second evaluation metric, the area under the prescriptive effect curve (AUPEC). Like the area under the receiver and operating characteristic curve (AUROC) for classification, the AUPEC represents an overall summary measure of how well an ITR performs over the random treatment rule that treats the same proportion of units.

We estimate these evaluation metrics using Neyman (1923)'s repeated sampling framework (see Imbens and Rubin, 2015, Chapter 6). An advantage of this approach is that it does not require any modeling assumption or asymptotic approximation. As a result, we can evaluate a broad class of ITRs including those based on complex machine learning algorithms. We show how to estimate the PAPE and AUPEC with a minimal amount of finite sample bias and derive the exact variance solely based on random sampling of units and random assignment of treatment.

We further extend this methodology to a common evaluation setting, in which the same experimental data is used to both estimate and evaluate ITRs. In this case, our finite-sample variance calculation is exact and directly incorporates the additional uncertainty due to random splits of data used for cross-validation. We implement the proposed methodology through an open-source R package.

Our simulation study demonstrates the accurate coverage of the proposed confidence intervals in small samples (Section 5). We also apply our methods to the Project STAR (Student-Teacher Achievement Ratio) experiment and compare the empirical performance of ITRs based on several popular methods (Section 6). Our evaluation approach addresses theoretical and practical difficulties of conducting reliable statistical inference for ITRs.

**Relevant Literature.** A large number of existing studies have focused on the derivation of optimal ITRs that maximize the population average value. For example, Qian and Murphy (2011) use penalized least squares whereas Zhao *et al.* (2012) show how a support vector machine can be used to derive an optimal ITR. Another popular approach is based on doubly-robust estimation (e.g., Dudík *et al.*, 2011; Zhang *et al.*, 2012; Chakraborty *et al.*, 2014; Jiang and Li, 2016; Athey and Wager, 2018; Kallus, 2018).

We propose a general methodology for empirically evaluating and comparing the performance of various ITRs including the ones proposed by these and other authors. While many of these methods come with uncertainty measures, even those that produce standard errors rely on asymptotic approximation, modeling assumptions, or resampling methods. In contrast, our methodology utilizes Neyman's repeated sampling framework and does not require any of these assumptions or approximations.

There also exists a related literature on policy evaluation. Starting with Manski (2004), many studies focus on the derivation of regret bounds given a class of ITRs. For example, Kitagawa and Tetenov (2018) show that an ITR, which maximizes the empirical average value, is minimax optimal without a strong restriction on the class of ITRs, whereas Athey and Wager (2018) establish a regret bound for an ITR based on doubly-robust estimation in observational studies (see also Zhou *et al.*, 2018). In addition, Luedtke and van der Laan (2016a,b) propose consistent estimators of the optimal average value even when an optimal ITR is not unique (see also Rai, 2018).

Our goal is different from these studies. We focus on statistical inference using the Neyman's repeated sampling framework for the experimental evaluation of arbitrary ITRs including optimal or non-optimal and simple or complex ones. Our evaluation metric is also different from the existing metrics. In particular, to the best of our knowledge, we are the first to formally study the AUPEC as an AUROC-like summary measure for evaluation.

In contrast, much of the policy evaluation literature focus on the optimal average value, which is required to compute the regret of an ITR. Athey and Wager (2018) briefly discusses a quantity related to the PAPE in their empirical application, but this quantity evaluates an ITR against the treatment rule that randomly treats exactly one half of units rather than the same proportion as the one treated under the ITR. Although empirical studies in the campaign and marketing literatures have used "uplift modeling," which is based on the PAPE (e.g., Imai and Strauss, 2011; Rzepakowski and Jaroszewicz, 2012; Gutierrez and Gérardy, 2016; Ascarza, 2018; Fifield, 2018), none develops formal estimation and inferential methods. We show that the AUPEC is a generalization of the QINI coefficient, which is a widely utilized statistic in uplift modeling (Radcliffe, 2007; Diemert *et al.*, 2018). Thus, our theoretical results for the AUPEC apply directly to the QINI coefficient as well.

Another related literature is concerned with the estimation of heterogeneous effects. Researchers have explored the use of tree-based methods (e.g., Imai and Strauss, 2011; Athey and Imbens, 2016; Wager and Athey, 2018; Hahn *et al.*, 2020), regularized regressions (e.g., Imai and Ratkovic, 2013; Künzel *et al.*, 2018), and ensamble methods (e.g., van der Laan and Rose, 2011; Grimmer *et al.*, 2017). In practice, the estimated heterogeneous treatment effects based on these machine learning algorithms are used to construct ITRs.

However, as Chernozhukov *et al.* (2019) point out, most machine learning algorithms, which require data-driven tuning parameters, cannot be regarded as consistent estimators of the conditional average treatment effect (CATE) unless strong assumptions are imposed. They propose a methodology to estimate heterogeneous treatment effects without such assumptions. Similar to theirs, our methodology does not depend on any modeling assumption and accounts for the uncertainty due to splitting of data. The key difference is that we focus on the evaluation of ITRs. In addition, our variance calculation is based on randomization and does not rely on asymptotic approximation.

Finally, Andrews *et al.* (2020) develops a conditional inference procedure, based on normal approximation, for the average value of the best-performing policy based on experimental or observational data. In contrast, we develop an unconditional exact inference for the difference in the average value between any pair of policies under a budget constraint. We also consider the evaluation of estimated policies based on the same experimental data using cross-validation whereas Andrews *et al.* focus on the evaluation of fixed policies.

# 2 Evaluation Metrics

In this section, we introduce our evaluation metrics. We first propose the population average prescriptive effect (PAPE) that, unlike the population average value, adjusts for the proportion of units treated by an ITR. The idea is that an efficacious ITR should outperform a non-individualized treatment rule, which randomly assigns the same proportion of units to the treatment condition. We extend the PAPE to the settings with a binding budget constraint. Finally, we propose the area under the prescriptive effect curve (AUPEC) as a univariate summary performance measure of an ITR under a range of budget constraint.

## 2.1 The Setup

Following the literature, we define an ITR as a deterministic map from the covariate space $\mathcal{X}$ to the binary treatment assignment (e.g., Qian and Murphy, 2011; Zhao *et al.*, 2012),

$$f : \mathcal{X} \longrightarrow \{0, 1\}.$$

Let $T_i$ denote the treatment assignment indicator variable, which is equal to $1$ if unit $i$ is assigned to the treatment condition, i.e., $T_i \in \mathcal{T} = \{0, 1\}$. For each unit, we observe the outcome variable $Y_i \in \mathcal{Y}$ as well as the vector of pre-treatment covariates, $\mathbf{X}_i \in \mathcal{X}$, where $\mathcal{Y}$ is the support of the outcome variable. We assume no interference between units and denote the potential outcome for unit $i$ under the treatment condition $T_i = t$ as $Y_i(t)$ for $t = 0, 1$. Then, the observed outcome is given by $Y_i = Y_i(T_i)$.

ASSUMPTION 1 (NO INTERFERENCE BETWEEN UNITS) *The potential outcomes for unit $i$ do not depend on the treatment status of other units. That is, for all $t_1, t_2, \ldots, t_n \in \{0, 1\}$, we have, $Y_i(T_1 = t_1, T_2 = t_2, \ldots, T_n = t_n) = Y_i(T_i = t_i)$.*

Under this assumption, the existing literature almost exclusively focuses on the derivation of an optimal ITR that maximizes the following population average value (e.g., Qian and Murphy, 2011; Zhao *et al.*, 2012; Zhou *et al.*, 2017),

$$\lambda_f \;=\; \mathbb{E}\{Y_i(f(\mathbf{X}_i))\}. \tag{1}$$

Next, we show that $\lambda_f$ may not be the best evaluation metric in some cases.

## 2.2 The Population Average Prescriptive Effect

We now introduce our main evaluation metric, the Population Average Prescriptive Effect (PAPE). The PAPE is based on two ideas. First, it is reasonable to expect a good ITR to outperform a *non-individualized* treatment rule, which does not use any information about

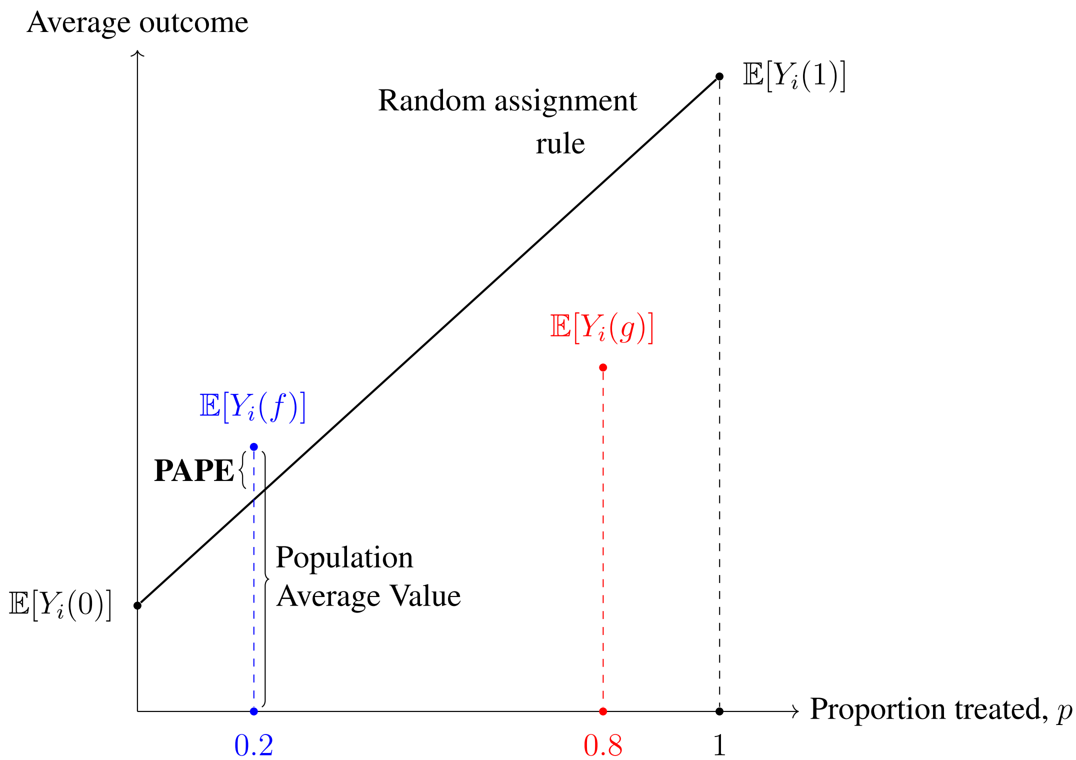


Figure 1: The Importance of Accounting for the Proportion of Treated Units. In this illustrative example, an ITR $g$ (red) outperforms another ITR $f$ (blue) in terms of the population average value, i.e., $\mathbb{E}[Y_i(f)] < \mathbb{E}[Y_i(g)]$. However, unlike $f$, the ITR $g$ is doing worse than the random treatment rule (black). In contrast, the population average prescriptive effect (PAPE) measures the performance of an ITR as the difference in the average value between the ITR and random treatment rule.

individual units when deciding who should receive the treatment. Second, a budget constraint should be considered since the treatment is often costly. This means that a good ITR should identify units who benefit from the treatment most. These two considerations lead to the random treatment rule, which assigns, with equal probability, the same proportion of units to the treatment condition, as a natural baseline for comparison.

Figure 1 illustrates the importance of accounting for the proportion of units treated by an ITR. In this figure, the horizontal axis represents the proportion treated and the vertical axis represents the average outcome under an ITR. The example shows that an ITR $g$ (red), which treats 80% of units, has a greater average value than another ITR $f$ (blue), which treats 20% of units, i.e., $\mathbb{E}[Y_i(f)] < \mathbb{E}[Y_i(g)]$. Despite this fact, $g$ is outperformed by the random treatment rule (black), which treats the same proportion of units, whereas $f$ does a better job than the random treatment rule. This is indicated by the fact that the black solid line is placed above $\mathbb{E}[Y_i(g)]$ and below $\mathbb{E}[Y_i(f)]$.

To overcome this undesirable property of the average value, we propose an alternative evaluation metric that compares the performance of an ITR with that of the random treatment rule. The random treatment rule serves as a natural baseline because it treats the same proportion of units without any individual information. This is analogous to the predictive setting, in which a classification algorithm is often compared to random classification.

Formally, let $p_f = \Pr(f(\mathbf{X}_i) = 1)$ denote the population proportion of units assigned to the treatment condition under ITR $f$. Without loss of generality, we assume a positive average treatment effect $\tau = \mathbb{E}\{Y_i(1) - Y_i(0)\} > 0$ so that the random treatment rule assigns the exactly proportion $p_f > 0$ of the units to the treatment group. If the treatment is on average harmful (a testable condition using the experimental data), the best random treatment rule is to treat no one. In that case, the estimation of the average value is sufficient for the evaluation. We define the population average prescription effect (PAPE) of ITR $f$ as the following difference in the average value between the ITR and random treatment rule,

$$\tau_f \;=\; \mathbb{E}\{Y_i(f(\mathbf{X}_i)) - p_f Y_i(1) - (1 - p_f)Y_i(0)\}. \tag{2}$$

One motivation for the PAPE is that administering a treatment is often expensive. Consider a costly treatment that does not harm anyone but only benefits a relatively small fraction of people. If we do not impose a budget constraint, treating everyone is the best ITR but such a policy does not use any individual-level information. Thus, to further evaluate the efficacy of an ITR, we extend the PAPE to the settings with a budget constraint.

## 2.3 Incorporating a Budget Constraint

With a budget constraint, we cannot simply treat all units who are predicted to benefit from the treatment. Instead, an ITR must be based on a *scoring rule* that sorts units according to their treatment priority: a unit with a greater score has a higher priority to receive the treatment. Let $s : \mathcal{X} \longrightarrow \mathcal{S}$ be such a scoring rule where $\mathcal{S} \subset \mathbb{R}$. For simplicity, we assume that the scoring rule is bijective, i.e., $s(\mathbf{x}) \neq s(\mathbf{x}')$ for any $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ with $\mathbf{x} \neq \mathbf{x}'$. This assumption is not restrictive as we can always redefine $\mathcal{X}$ such that the assumption holds.

We define an ITR based on a scoring rule by assigning a unit to the treatment group if and only if its score is higher than a threshold, $c$,

$$f(\mathbf{X}_i, c) \ = \ \mathbf{1}\{s(\mathbf{X}_i) > c\}.$$

Under a binding budget constraint $p$, we define the threshold that corresponds to the maximal proportion of treated units under the budget constraint, i.e.,

$$c_p(f) \ = \ \inf\{c \in \mathbb{R} : \Pr(f(\mathbf{X}_i, c) = 1) \leq p\}.$$

Our framework allows for any arbitrary scoring rule. A popular scoring rule is the conditional average treatment effect (CATE),

$$s(\mathbf{x}) \ = \ \mathbb{E}(Y_i(1) - Y_i(0) \mid \mathbf{X}_i = \mathbf{x}).$$

Researchers have studied the estimation of the CATE using various machine learning algorithms such as tree-based methods and regularized regressions.

We emphasize that the scoring rule need not be based on the CATE. In fact, policy makers rely on various indexes. Such examples include the MELD (Kamath *et al.*, 2001) that determines liver transplant priority, and the IDA Resource Allocation Index that informs the World Bank about the provision of economic aid.

Given this setup, we generalize the PAPE to the setting with a budget constraint. As before, without loss of generality, we assume the treatment is on average beneficial, i.e., $\tau = \mathbb{E}\{Y_i(1) - Y_i(0)\} > 0$, so that the constraint is binding for the random treatment rule treating at most $100 \times p\%$ of units. The PAPE with a budget constraint $p$ is defined as,

$$\tau_{fp} \ = \ \mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p(f))) - pY_i(1) - (1 - p)Y_i(0)\}. \tag{3}$$

A budget constraint facilitates the comparison of multiple ITRs on the same footing. Suppose that we compare two ITRs, $f$ and $g$, using the difference in their average values,

$$\Delta(f,g) \;=\; \lambda_f - \lambda_g \;=\; \mathbb{E}\{Y_i(f(\mathbf{X}_i)) - Y_i(g(\mathbf{X}_i))\}. \tag{4}$$

While this quantity is useful, like the average value, it also fails to take into account the proportion of units assigned to the treatment condition under each ITR.

We can address this issue by comparing the efficacy of two ITRs under the same budget constraint. Formally, we define the Population Average Prescriptive Effect Difference (PAPD) under budget $p$ as,

$$\Delta_p(f,g) \;=\; \tau_{fp} - \tau_{gp} \;=\; \mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p(f))) - Y_i(g(\mathbf{X}_i, c_p(g)))\}. \tag{5}$$

## 2.4 The Area under the Prescriptive Effect Curve

Since the PAPE (Eqn (3)) varies as a function of budget constraint $p$, it would be useful to develop a summary performance metric of an ITR over a range of $p$. We propose the area under the prescriptive effect curve (AUPEC) as a metric analogous to the area under the receiver operating characteristic curve (AUROC) for classification performance.

Figure 2 graphically illustrates the AUPEC. Similar to Figure 1, the vertical and horizontal axes represent the average outcome and the budget, respectively. The budget is operationalized as the maximal proportion treated. The red solid curve corresponds to the average value of an ITR $f$ as a function of budget constraint $p$, i.e., $\mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p(f)))\}$, whereas the black solid line represents the average value of the random treatment rule. The AUPEC corresponds to the area under the red curve minus the area under the black line, which is shown as a red shaded area.

Thus, the AUPEC represents the average performance of an ITR relative to the random treatment rule over the entire range of budget constraint (one could also compute the AUPEC over a specific range of budgets). Unlike the previous work (e.g., Rzepakowski and Jaroszewicz, 2012), we do not require an ITR to assign the maximal proportion of units

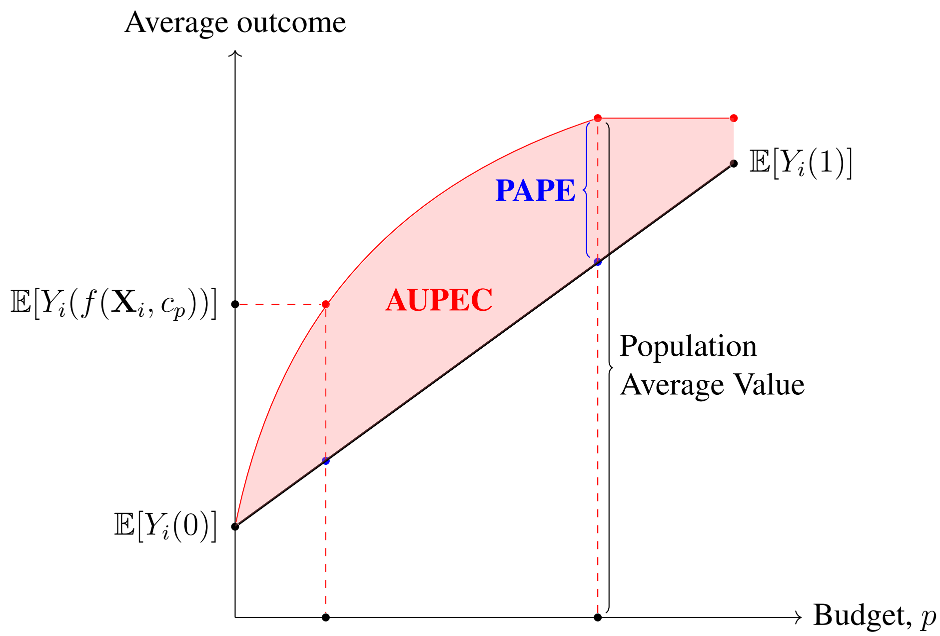


Figure 2: The Area Under the Prescriptive Effect Curve (AUPEC). The black solid line represents the average value under the random treatment rule while the red solid curve represents the average value under an ITR $f$. The difference between the line and the curve at a given budget constraint corresponds to the population average prescriptive effect (PAPE). The shaded area between the line and the curve represents the AUPEC of $f$.

to the treatment condition though such a constraint can also be imposed if desired. For example, treating more than a certain proportion of units will reduce the average outcome if these additional units are harmed by the treatment. This is indicated by the flatness of the red line after $p_f$ in Figure 2.

Formally, for a given ITR $f$, we define the AUPEC as,

$$\Gamma_f = \int_0^{p_f} \mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p(f)))\} dp + (1 - p_f)\mathbb{E}\{Y_i(f(\mathbf{X}_i, c^*))\} - \frac{1}{2}\mathbb{E}(Y_i(0) + Y_i(1)), \quad (6)$$

where $c^*$ is the pre-determined minimum score such that one would be treated in the absence of a budget constraint, and $p_f = \Pr(f(\mathbf{X}_i, c^*) = 1)$ denotes the maximal proportion of units assigned to the treatment condition under the ITR with no budget constraint. The last term represents the area under the random treatment rule.

Different values of the threshold $c^*$ are possible. If the goal is to treat only those who on average benefit from the treatment, then we could use the CATE as a scoring rule and set $c^* = 0$. Another example is the use of the MELD score as the scoring rule and choose an appropriate value of $c^*$ so that a sufficiently healthy patient is never considered for a transplant. Finally, setting $c^* = -\infty$ would represent the setting where the maximum possible units should be treated regardless of the scoring rule.

We further note that the AUPEC is a generalization of the QINI coefficient widely utilized in literature for uplift modeling (Radcliffe, 2007). Formally, the population-level QINI coefficient is commonly defined as,

$$\text{QINI} \;=\; n\left(\int_0^1 p\mathbb{E}\{Y_i(1) - Y_i(0) \mid f(\mathbf{X}_i, c_p(f)) = 1\}\mathrm{d}p - \frac{1}{2}\mathbb{E}(Y_i(1) - Y_i(0))\right).$$

After some algebra, we can rewrite this quantity in the following form,

$$n\left(\int_0^1 \mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p(f)))\}\mathrm{d}p - \frac{1}{2}\mathbb{E}(Y_i(0) + Y_i(1))\right)$$

Thus, the QINI coefficient is (up to a constant factor $n$) a special case of AUPEC when $c^* = -\infty$ and $p_f = 1$. The choice of $c^* = -\infty$ may be reasonable in the applications where the treatment can be assumed to be never harmful, i.e., $Y_i(1) \geq Y_i(0)$. In such cases, under no budget constraint one would treat the entire population.

To enable a comparison of efficacy across different datasets, the AUPEC can be normalized to be scale-invariant by shifting $\Gamma_f$ by $\mathbb{E}(Y_i(0))$ and dividing by $\tau = \mathbb{E}(Y_i(1) - Y_i(0))$,

$$\Gamma_f^* \;=\; \frac{1}{\tau}\left[\int_0^{p_f} \mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p(f)))\}\mathrm{d}p + (1 - p_f)\mathbb{E}\{Y_i(f(\mathbf{X}_i, c^*))\} - \mathbb{E}(Y_i(0))\right] - \frac{1}{2}.$$

This normalized AUPEC is invariant to the affine transformation of the outcome variables, $Y_i(1), Y_i(0)$, while the standard AUPEC is only invariant to a constant shift. The normalized AUPEC takes a value in $[0, 1]$, and has an intuitive interpretation as the average percentage outcome gain using the ITR $f$ compared to the random treatment rule, under the uniform prior distribution over the percentage treated.

# 3 Estimation and Inference

Having introduced the evaluation metrics, we show how to estimate them and compute standard errors under the repeated sampling framework of Neyman (1923). Here, we consider the setting, in which researchers are interested in evaluating the performance of fixed ITRs. In other words, throughout this section, ITR $f$ is assumed to be known and has no estimation uncertainty. For example, one may first construct an ITR based on an existing data set (experimental or observational), and then conduct a new experiment to evaluate its performance. In Section 4, we extend the methodology developed here to the setting, in which the same experimental data set is used to both construct and evaluate an ITR.

For the rest of this paper, we assume that we have a simple random sample of $n$ units from a super-population, $\mathcal{P}$. We conduct a completely randomized experiment, in which $n_1$ units are randomly assigned to the treatment condition with probability $n_1/n$ and the rest of the $n_0 (= n - n_1)$ units are assigned to the control condition. While it is straightforward to allow for unequal treatment assignment probabilities across units, for the sake of simplicity, we assume complete randomization. We formalize these assumptions below.

ASSUMPTION 2 (RANDOM SAMPLING OF UNITS) *Each of $n$ units, represented by a three-tuple consisting of two potential outcomes and pre-treatment covariates, is assumed to be independently sampled from a super-population $\mathcal{P}$, i.e.,*

$$(Y_i(1), Y_i(0), \mathbf{X}_i) \overset{\text{i.i.d.}}{\sim} \mathcal{P}$$

ASSUMPTION 3 (COMPLETE RANDOMIZATION) *For any $i = 1, 2, \ldots, n$, the treatment assignment probability is given by,*

$$\Pr(T_i = 1 \mid Y_i(1), Y_i(0), \mathbf{X}_i) = \frac{n_1}{n}$$

*where* $\sum_{i=1}^{n} T_i = n_1$.

We now present the results under a binding budget constraint. The results for the average value and PAPE with no budget constraint appear in Appendix A.1.

## 3.1 The Population Average Prescription Effect (PAPE)

To estimate the PAPE with a binding budget constraint $p$ (Eqn (3)), we consider the following estimator,

$$\hat{\tau}_{fp}(\mathcal{Z}) = \frac{1}{n_1}\sum_{i=1}^{n} Y_i T_i f(\mathbf{X}_i, \hat{c}_p(f)) + \frac{1}{n_0}\sum_{i=1}^{n} Y_i (1-T_i)(1-f(\mathbf{X}_i, \hat{c}_p(f))) - \frac{p}{n_1}\sum_{i=1}^{n} Y_i T_i - \frac{1-p}{n_0}\sum_{i=1}^{n} Y_i(1-T_i) \tag{7}$$

where $\hat{c}_p(f) = \inf\{c \in \mathbb{R} : \sum_{i=1}^n f(\mathbf{X}_i, c) \leq np\}$ represents the estimated threshold given the maximal proportion of treated units $p$. Unlike the case of no budget constraint (see Appendix A.1.2), the bias is not zero because the threshold $c_p(f)$ needs to be estimated without an assumption about the distribution of the score produced by the scoring rule. We derive an upper bound of bias and the exact variance.

THEOREM 1 (BIAS BOUND AND EXACT VARIANCE OF THE PAPE ESTIMATOR WITH A BUDGET CONSTRAINT) *Under Assumptions 1, 2, and 3, the bias of the proposed estimator of the PAPE with a budget constraint* $p$ *defined in Eqn* (7) *can be bounded as follows,*

$$\begin{aligned} & \mathbb{P}_{\hat{c}_p(f)}(|\mathbb{E}\{\hat{\tau}_{fp}(\mathcal{Z}) - \tau_{fp} \mid \hat{c}_p(f)\}| \geq \epsilon) \\ \leq \;& 1 - B(1-p+\gamma_{fp}(\epsilon), n - \lfloor np \rfloor, \lfloor np \rfloor + 1) + B(1-p-\gamma_{fp}(\epsilon), n - \lfloor np \rfloor, \lfloor np \rfloor + 1), \end{aligned}$$

*where any given constant* $\epsilon > 0$, $B(\epsilon, \alpha, \beta)$ *is the incomplete beta function (if* $\alpha \leq 0$ *and* $\beta > 0$, *we set* $B(\epsilon, \alpha, \beta) := H(\epsilon)$ *for all* $\epsilon$ *where* $H(\epsilon)$ *is the Heaviside step function), and*

$$\gamma_{fp}(\epsilon) = \frac{\epsilon}{\max_{c \in [c_p(f)-\epsilon,\, c_p(f)+\epsilon]} \mathbb{E}(\tau_i \mid s(\mathbf{X}_i) = c)}.$$

*The variance of the estimator is given by,*

$$\mathbb{V}(\hat{\tau}_{fp}(\mathcal{Z})) = \frac{\mathbb{E}(S_{fp1}^2)}{n_1} + \frac{\mathbb{E}(S_{fp0}^2)}{n_0} + \frac{\lfloor np \rfloor (n - \lfloor np \rfloor)}{n^2(n-1)} \left\{(2p-1)\kappa_{f1}(p)^2 - 2p\kappa_{f1}(p)\kappa_{f0}(p)\right\},$$

*where* $S_{fpt}^2 = \sum_{i=1}^n (Y_{fpi}(t) - \overline{Y_{fp}(t)})^2/(n-1)$ *and* $\kappa_{ft}(p) = \mathbb{E}(\tau_i \mid f(\mathbf{X}_i, \hat{c}_p(f)) = t)$ *with* $Y_{fpi}(t) = (f(\mathbf{X}_i, \hat{c}_p(f)) - p)\, Y_i(t)$, *and* $\overline{Y_{fp}(t)} = \sum_{i=1}^n Y_{fpi}(t)/n$, *for* $t = 0, 1$.

Proof is given in Appendix A.2. The last term in the variance accounts for the variance due to estimating $c_p(f)$. The variance can be consistently estimated by replacing each unknown parameter with its sample analogue, i.e., for $t = 0, 1$,

$$\begin{aligned} \widehat{\mathbb{E}(S_{fpt}^2)} &= \frac{1}{n_t - 1} \sum_{i=1}^n \mathbf{1}\{T_i = t\}(Y_{fpi} - \overline{Y_{fpt}})^2, \\ \widehat{\kappa_{ft}(p)} &= \frac{\sum_{i=1}^n \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_p(f)) = t\} T_i Y_i}{\sum_{i=1}^n \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_p(f)) = t\} T_i} - \frac{\sum_{i=1}^n \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_p(f)) = t\}(1-T_i) Y_i}{\sum_{i=1}^n \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_p(f)) = t\}(1-T_i)}, \end{aligned}$$

where $Y_{fpi} = (f(\mathbf{X}_i, \hat{c}_p(f)) - p) Y_i$ and $\overline{Y}_{fpt} = \sum_{i=1}^n \mathbf{1}\{T_i = t\} Y_{fpi}/n_t$. To estimate the term that appears in the denominator of $\gamma_{fp}(\epsilon)$ as part of the upper bound of bias, we may assume that the CATE, $\mathbb{E}(\tau_i \mid s(\mathbf{X}_i) = c)$, is continuous in $c$. Continuity is often assumed when estimating the CATE (e.g., Künzel *et al.*, 2018; Wager and Athey, 2018). We may also utilize an upper bound for CATE, if known, to estimate the bias conservatively.

Building on the above results, we also consider the comparison of two ITRs under the same budget constraint, using the following estimator of the PAPD (Eqn (5)),

$$\widehat{\Delta}_p(f,g,\mathcal{Z}) \;=\; \frac{1}{n_1}\sum_{i=1}^{n} Y_iT_i\{f(\mathbf{X}_i,\hat{c}_p(f))-g(\mathbf{X}_i,\hat{c}_p(g))\}+\frac{1}{n_0}\sum_{i=1}^{n} Y_i(1-T_i)\{g(\mathbf{X}_i,\hat{c}_p(g))-f(\mathbf{X}_i,\hat{c}_p(f))\}. \tag{8}$$

Theorem A3 in Appendix A.3 derives the bias bound and exact variance of this estimator. In one of their applications, Zhou *et al.* (2018) apply the $t$-test to the cross-validated test statistic similar to the one introduced here under no budget constraint. However, no formal justification of this procedure is given under the cross-validation setting and it cannot be readily extended to the case with a budget constraint. In contrast, our methodology is applicable under these settings as well (see Section 4).

## 3.2 The Area under the Prescriptive Effect Curve

Next, we consider the estimation and inference about the AUPEC (Eqn (6)). Let $n_f$ represent the maximum number of units in the sample that the ITR $f$ would assign under no budget constraint, i.e., $\hat{p}_f = n_f/n = \sum_{i=1}^n f(\mathbf{X}_i, c^*)/n$. We propose the following estimator of the AUPEC,

$$\begin{aligned}\widehat{\Gamma}_f(\mathcal{Z}) \;=\; & \frac{1}{n_1}\sum_{i=1}^{n} Y_iT_i\left\{\frac{1}{n}\left(\sum_{k=1}^{n_f} f(\mathbf{X}_i,\hat{c}_{k/n}(f))+(n-n_f)f(\mathbf{X}_i,\hat{c}_{\hat{p}_f}(f))\right)\right\}\\ & +\frac{1}{n_0}\sum_{i=1}^{n} Y_i(1-T_i)\left\{1-\frac{1}{n}\left(\sum_{k=1}^{n_f} f(\mathbf{X}_i,\hat{c}_{k/n}(f))+(n-n_f)f(\mathbf{X}_i,\hat{c}_{\hat{p}_f}(f))\right)\right\}\\ & -\frac{1}{2n_1}\sum_{i=1}^{n} Y_iT_i-\frac{1}{2n_0}\sum_{i=1}^{n} Y_i(1-T_i). \end{aligned} \tag{9}$$

The following theorem shows a bias bound and the exact variance of this estimator.

THEOREM 2 (BIAS AND VARIANCE OF THE AUPEC ESTIMATOR) *Under Assumptions 1, 2, and 3, the bias of the AUPEC estimator defined in Eqn* (9) *can be bounded as follows,*

$$\begin{aligned}\mathbb{P}_{\hat{p}_f}(|\mathbb{E}(\widehat{\Gamma}_f(\mathcal{Z})-\Gamma_f \mid \hat{p}_f)| \geq \epsilon) \;\leq\; & 1-B(1-p_f+\gamma_{p_f}(\epsilon), n-\lfloor np_f\rfloor, \lfloor np_f\rfloor+1)\\ & + B(1-p_f-\gamma_{p_f}(\epsilon), n-\lfloor np_f\rfloor, \lfloor np_f\rfloor+1)\end{aligned}$$

*where any given constant* $\epsilon > 0$, $B(\epsilon,\alpha,\beta)$ *is the incomplete beta function (if* $\alpha = 0$ *and* $\beta > 0$, *we set* $B(\epsilon,\alpha,\beta) := H(\epsilon)$ *for all* $\epsilon$ *where* $H(\epsilon)$ *is the Heaviside step function), and*

$$\gamma_{p_f}(\epsilon) \;=\; \frac{\epsilon}{2\max_{c\in[c^*-\epsilon,\; c^*+\epsilon]}\mathbb{E}(\tau_i \mid s(\mathbf{X}_i)=c)}.$$

*The variance is given by,*

$$
\begin{aligned}
& \mathbb{V}(\widehat{\Gamma}_f(\mathcal{Z})) \\
= \quad & \frac{\mathbb{E}(S_{f1}^{*2})}{n_1} + \frac{\mathbb{E}(S_{f0}^{*2})}{n_0} + \mathbb{E}\left[ -\frac{1}{n} \left\{ \sum_{z=1}^{Z} \frac{z(n-z)}{n^2(n-1)} \kappa_{f1}(z/n)\kappa_{f0}(z/n) + \frac{Z(n-Z)^2}{n^2(n-1)} \kappa_{f1}(Z/n)\kappa_{f0}(Z/n) \right\} \right. \\
& \quad - \frac{2}{n^4(n-1)} \sum_{z=1}^{Z-1} \sum_{z'=z+1}^{Z} z(n-z')\kappa_{f1}(z/n)\kappa_{f1}(z'/n) - \frac{Z^2(n-Z)^2}{n^4(n-1)} \kappa_{f1}(Z/n)^2 \\
& \quad \left. - \frac{2(n-Z)^2}{n^4(n-1)} \sum_{z=1}^{Z} z\kappa_{f1}(Z/n)\kappa_{f1}(z/n) + \frac{1}{n^4} \sum_{z=1}^{Z} z(n-z)\kappa_{f1}(z/n)^2 \right] \\
& + \mathbb{V}\left( \sum_{z=1}^{Z} \frac{z}{n} \kappa_{f1}(z/n) + \frac{(n-Z)Z}{n} \kappa_{f1}(Z/n) \right),
\end{aligned}
$$

*where $Z$ is a Binomial random variable with size $n$ and success probability $p_f$, and $S_{ft}^{*2} = \sum_{i=1}^{n}(Y_i^*(t) - \overline{Y^*(t)})^2/(n-1)$, $\kappa_{ft}(k/n) = \mathbb{E}(Y_i(1) - Y_i(0) \mid f(\mathbf{X}_i, \hat{c}_{k/n}(f)) = t)$, with $Y_i^*(t) = \left[ \left\{ \sum_{z=1}^{n_f} f(\mathbf{X}_i, \hat{c}_{z/n}(f)) + (n - n_f) f(\mathbf{X}_i, \hat{c}_{\hat{p}_f}(f)) \right\} / n - \frac{1}{2} \right] Y_i(t)$ and $\overline{Y^*(t)} = \sum_{i=1}^{n} Y_i^*(t)/n$, for $t = 0, 1$.*

Proof is given in in Appendix A.4. When $c^* = -\infty$ (i.e., the AUPEC equals the QINI coefficient), the estimator is unbiased, implying that the bias comes from estimating the proportion treated $p_f$ under no budget constraints and $c^* > -\infty$. As before, $\mathbb{E}({S_{ft}^*}^2)$ does not equal $\mathbb{V}(Y_i^*(t))$ due to the need to estimate the terms $c_{z/n}$ for all $z$, and the additional terms account for the variance of estimation. We can consistently estimate the upper bound of bias, for example, under by assuming that the CATE is continuous. We may also utilize an upper bound for CATE, if known, to estimate the bias conservatively.

To estimate the variance, we replace each unknown parameter with its sample analogue,

$$
\begin{aligned}
\widehat{\mathbb{E}(S_{ft}^{*2})} &= \frac{1}{n_t - 1} \sum_{i=1}^{n} \mathbf{1}\{T_i = t\}(Y_i^* - \overline{Y_t^*})^2, \\
\hat{\kappa}_{ft}(z/n) &= \frac{\sum_{i=1}^{n} \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_{z/n}(f)) = t\} T_i Y_i}{\sum_{i=1}^{n} \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_{z/n}(f)) = t\} T_i} - \frac{\sum_{i=1}^{n} \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_{z/n}(f)) = t\}(1 - T_i) Y_i}{\sum_{i=1}^{n} \mathbf{1}\{f(\mathbf{X}_i, \hat{c}_{z/n}(f)) = t\}(1 - T_i)},
\end{aligned}
\tag{10}
$$

for $t = 0, 1$. In the extreme cases with $z \to 1$ for $t = 1$ and $z \to n$ for $t = 0$, each denominator in Eqn (10) is likely to be close to zero. In such cases, we instead use the estimator $\kappa_{f1}\widehat{(z_{\min}/n)}$ for all $z < z_{\min}$ where $z_{\min}$ is the smallest $z$ such that Eqn (10) for

$\kappa_{f1}(z/n)$ does not lead to division by zero. Similarly, for $t = 0$, we use $\kappa_{f0}\widehat{(z_{\max}/n)}$ for all $z > z_{\max}$ where $z_{\max}$ is the largest $z$.

For the terms involving the binomial random variable $Z$, we first note that, when fully expanded out, they are the polynomials of $p_f = \mathbb{E}(f(\mathbf{X}_i))$. To estimate the polynomials, we can utilize their unbiased estimators as discussed in Stuard and Ord (1994), i.e., $\hat{p}_f^z = s(s-1)\cdots(s-z+1)/\{n(n-1)\cdots(n-z+1)\}$ where $s = \sum_{i=1}^n f(\mathbf{X}_i)$ is unbiased for $p_f^z$ for all $z \leq n$. When the sample size is large, this estimation method is computationally inefficient and unstable due to the presence of high powers. Hence, we may use the Monte Carlo sampling of $Z$ from a Binomial distribution with size $n$ and success probability $\hat{p}_f$. In our simulation study, we show that this Monte Carlo approach is effective even when the sample size is small (see Section 5).

Finally, a consistent estimator for the normalized AUPEC is given by,

$$\begin{aligned}
\widehat{\Gamma}_f^*(\mathcal{Z}) \quad = \quad & \frac{1}{\sum_{i=1}^n Y_i T_i/n_1 - Y_i(1-T_i)/n_0} \left\{ \frac{1}{n n_1} \sum_{i=1}^n Y_i T_i \left( \sum_{z=1}^{n_f} f(\mathbf{X}_i, \hat{c}_{z/n}(f)) + (n - n_f) f(\mathbf{X}_i, \hat{c}_{\hat{p}_f}(f)) \right) \right. \\
& \left. - \frac{1}{n n_0} \sum_{i=1}^n Y_i (1 - T_i) \left( \sum_{z=1}^{n_f} f(\mathbf{X}_i, \hat{c}_{z/n}(f)) + (n - n_f) f(\mathbf{X}_i, \hat{c}_{\hat{p}_f}(f)) \right) \right\} - \frac{1}{2}. \qquad (11)
\end{aligned}$$

The variance of $\widehat{\Gamma}_f^*(\mathcal{Z})$ can be estimated using the Taylor expansion of quotients of random variables to an appropriate order as detailed in Stuard and Ord (1994).

# 4 Estimating and Evaluating ITRs Using the Same Experimental Data

We next consider a common situation, in which researchers use the same experimental data to both estimate and evaluate an ITR via cross-validation. This differs from the setting we have analyzed so far, in which a fixed ITR is given for evaluation. We first extend the evaluation metrics introduced in Section 2 to the current setting with estimated ITRs. We then develop inferential methods under the Neyman's repeated sampling framework by accounting for both estimation and evaluation uncertainties. Below, we consider the scenario, in which researchers face a binding budget constraint. Appendix A.5 presents the results for the case with no budget constraint.

## 4.1 Evaluation Metrics

Suppose that we have the data from a completely randomized experiment as described in Section 3. We first estimate an ITR $f$ by applying a machine learning algorithm $F$ to training data $\mathcal{Z}^{tr}$. Then, under a budget constraint of the maximal proportion of treated units $p$, we use test data to evaluate the resulting estimated ITR $\hat{f}_{\mathcal{Z}^{tr}}$. As before, we assume that this constraint is binding, i.e., $p < p_F$ where $p_F = \Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = 1)$ represents the proportion of treated units under the ITR without a budget constraint.

Formally, an machine learning algorithm $F$ is a deterministic map from the space of training data $\mathcal{Z}$ to that of scoring rules $\mathcal{S}$,

$$F : \mathcal{Z} \to \mathcal{S}.$$

Then, for a given training data set $\mathcal{Z}^{tr}$, the estimated ITR is given by,

$$\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c_p(\hat{f}_{\mathcal{Z}^{tr}})) \ = \ \mathbf{1}\{\hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) > c_p(\hat{f}_{\mathcal{Z}^{tr}})\},$$

where $\hat{s}_{\mathcal{Z}^{tr}} = F(\mathcal{Z}^{tr})$ is the estimated scoring rule and $c_p(\hat{f}_{\mathcal{Z}^{tr}}) \ = \ \inf\{c \in \mathbb{R} : \Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c) = 1 \mid \mathcal{Z}^{tr}) \leq p\}$ is the threshold based on the maximal proportion of treated units $p$. The CATE is a natural choice for the scoring rule, i.e., $s_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = \mathbb{E}(\tau_i \mid \mathbf{X}_i)$. We need not assume that $\hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)$ is consistent for the CATE.

To extend the PAPE (Eqn (3)), we first define the population proportion of units with $\mathbf{X}_i = \mathbf{x}$ who are assigned to the treatment condition under the estimated ITR as,

$$\bar{f}_{Fp}(\mathbf{x}) \ = \ \mathbb{E}_{\mathcal{Z}^{tr}}\{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c_p(\hat{f}_{\mathcal{Z}^{tr}})) \mid \mathbf{X}_i = \mathbf{x}\} \ = \ \Pr\{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c_p(\hat{f}_{\mathcal{Z}^{tr}})) = 1 \mid \mathbf{X}_i = \mathbf{x}\}.$$

While $c_p(\hat{f}_{\mathcal{Z}^{tr}})$ depends on the specific ITR generated from the training data $\mathcal{Z}^{tr}$, the population proportion of treated units averaged over the sampling of training data, $\bar{f}_{Fp}(\mathbf{X}_i)$, only depends on $p$.

Lastly, the PAPE of the estimated ITR under budget constraint $p$ is defined as,

$$\tau_{Fp} \;=\; \mathbb{E}\{\bar{f}_{Fp}(\mathbf{X}_i)Y_i(1) + (1-\bar{f}_{Fp}(\mathbf{X}_i))Y_i(0) - pY_i(1) - (1-p)Y_i(0)\}.$$

This evaluation metric corresponds to neither that of a specific ITR estimated from the whole experimental data set nor its expectation. Rather, we are evaluating the efficacy of a learning algorithm that is used to estimate an ITR using the same experimental data.

We can also compare *estimated* ITRs by further generalizing the definition of the PAPD (Eqn (5)) to the current setting. Specifically, we define the PAPD between two machine learning algorithms, $F$ and $G$, under budget constraint $p$ as,

$$\Delta_p(F,G) \;=\; \mathbb{E}_{\mathbf{X},Y}[\{\bar{f}_{Fp}(\mathbf{X}_i) - \bar{f}_{Gp}(\mathbf{X}_i)\}Y_i(1) + \{\bar{f}_{Gp}(\mathbf{X}_i) - \bar{f}_{Fp}(\mathbf{X}_i)\}Y_i(0)]. \qquad (12)$$

Finally, we consider the AUPEC of an estimated ITR. Specifically, the AUPEC of an machine learning algorithm $F$ is defined as,

$$\Gamma_F \;=\; \mathbb{E}_{\mathcal{Z}^{tr}}\left[\int_0^{p_{\hat{f}}} \mathbb{E}\{Y_i(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c_p(\hat{f}_{\mathcal{Z}^{tr}})))\mathrm{d}p + (1-p_{\hat{f}})\mathbb{E}\{Y_i(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c^*))\}\right] - \frac{1}{2}\mathbb{E}(Y_i(0)+Y_i(1)), \qquad (13)$$

where $p_{\hat{f}} = \Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = 1)$ is the maximal population proportion of units treated by the estimated ITR $\hat{f}_{\mathcal{Z}^{tr}}$.

## 4.2 Estimation and Inference

Rather than simply splitting the data into training and test sets (in such a case, the inferential procedure for fixed ITRs is applicable), we maximize efficiency by using cross-validation to estimate the evaluation metrics introduced above. First, we randomly split the data into $K$ subsamples of equal size $m = n/K$ by assuming, for the sake of notational simplicity, that $n$ is a multiple of $K$. Then, for each $k = 1, 2, \ldots, K$, we use the $k$th subsample as a test set $\mathcal{Z}_k = \{\mathbf{X}_i^{(k)}, T_i^{(k)}, Y_i^{(k)}\}_{i=1}^m$ with the data from all $(K-1)$ remaining subsamples as the training set $\mathcal{Z}_{-k} = \{\mathbf{X}_i^{(-k)}, T_i^{(-k)}, Y_i^{(-k)}\}_{i=1}^{n-m}$.

To simplify notation, we assume that the number of treated (control) units is identical across different folds and denote it as $m_1$ ($m_0 = m - m_1$). For each split $k$, we estimate an

**Algorithm 1** Estimating and Evaluating an Individualized Treatment Rule (ITR) using the Same Experimental Data via Cross-Validation

**Input**: Data $\mathcal{Z} = \{\mathbf{X}_i, T_i, Y_i\}_{i=1}^n$, Machine learning algorithm $F$, Evaluation metric $\tau_f$, Number of folds $K$

**Output**: Estimated evaluation metric $\hat{\tau}_F$, Estimated variance of $\hat{\tau}_F$

```
Split data into K random subsets of equal size (Z_1, ..., Z_k)
k ← 1
while k ≤ K do
    Z_{-k} = [Z_1, ..., Z_{k-1}, Z_{k+1}, ..., Z_K]
    f̂_{-k} = F(Z_{-k})                    ▷ Estimate ITR f by applying F to Z_{-k}
    τ̂_k = τ̂_{f̂_{-k}}(Z_k)                  ▷ Evaluate estimated ITR f̂ using Z_k
    k ← k + 1
end while
return τ̂_F = (1/K) Σ_{k=1}^K τ̂_k, V̂(τ̂_F) = v(f̂_{-1}, ..., f̂_{-k}, Z_1, ..., Z_K)
```

ITR by applying a learning algorithm $F$ to the training data $\mathcal{Z}_{-k}$,

$$\hat{f}_{-k} \; = \; F(\mathcal{Z}_{-k}). \tag{14}$$

We then evaluate the performance of the learning algorithm $F$ by computing an evaluation metric of interest $\tau$ based on the test data $\mathcal{Z}_k$. Repeating this $K$ times for each $k$ and averaging the results gives a cross-validation estimator of the evaluation metric. Algorithm 1 formally presents this estimation procedure. The variance formula for the estimated evaluation metric is omitted as it is generally a complex function $v(\cdot)$ (see Theorem 3 below).

We develop the inferential methodology for the evaluation based on the cross-validation procedure described above under the Neyman's repeated sampling framework. We focus on the case with a binding budget constraint. The results with no budget constraint appear in Appendix A.5. We begin by introducing the cross-validation estimator of the PAPE with a binding budget constraint $p$,

$$\hat{\tau}_{Fp} \; = \; \frac{1}{K}\sum_{k=1}^{K} \hat{\tau}_{\hat{f}_{-k},p}(\mathcal{Z}_k), \tag{15}$$

where $\hat{\tau}_{fp}$ is defined in Eqn (7).

Like the fixed ITR case, the bias of the proposed estimator is not exactly zero. However, we are able to show that the bias can be upper bounded by a small quantity while the exact randomization variance can still be derived.

THEOREM 3 (BIAS BOUND AND EXACT VARIANCE OF THE CROSS-VALIDATION PAPE ESTIMATOR WITH A BUDGET CONSTRAINT) *Under Assumptions 1, 2, and 3, the bias of the cross-validation PAPE estimator with a budget constraint* $p$ *defined in Eqn* (15) *can be bounded as follows,*

$$\mathbb{E}_{\mathcal{Z}^{tr}}[\mathbb{P}_{\hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})}(|\mathbb{E}\{\hat{\tau}_{Fp} - \tau_{Fp} \mid \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})\}| \geq \epsilon)] \;\leq\; 1 - B(1-p+\gamma_p(\epsilon), m - \lfloor mp \rfloor, \lfloor mp \rfloor + 1) + B(1-p-\gamma_p(\epsilon), m - \lfloor mp \rfloor, \lfloor mp \rfloor + 1),$$

*where any given constant* $\epsilon > 0$, $B(\epsilon, \alpha, \beta)$ *is the incomplete beta function (if* $\alpha = 0$ *and* $\beta > 0$, *we set* $B(\epsilon, \alpha, \beta) := H(\epsilon)$ *for all* $\epsilon$ *where* $H(\epsilon)$ *is the Heaviside step function), and*

$$\gamma_p(\epsilon) \;=\; \frac{\epsilon}{\mathbb{E}_{\mathcal{Z}^{tr}}\{\max_{c \in [c_p(\hat{f}_{\mathcal{Z}^{tr}})-\epsilon,\; c_p(\hat{f}_{\mathcal{Z}^{tr}})+\epsilon]} \mathbb{E}_{\mathcal{Z}}(\tau_i \mid \hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = c)\}}.$$

*The variance of the estimator is given by,*

$$\mathbb{V}(\hat{\tau}_{Fp}) \;=\; \frac{\mathbb{E}(S^2_{\hat{f}p1})}{m_1} + \frac{\mathbb{E}(S^2_{\hat{f}p0})}{m_0} + \frac{\lfloor mp \rfloor (m - \lfloor mp \rfloor)}{m^2(m-1)} \left\{(2p-1)\kappa_{F1}(p)^2 - 2p\kappa_{F1}(p)\kappa_{F0}(p)\right\} - \frac{K-1}{K}\mathbb{E}(S^2_{Fp}),$$

*where* $S^2_{\hat{f}pt} = \sum_{i=1}^{m}(Y_{\hat{f}pi}(t) - \overline{Y_{\hat{f}p}(t)})^2/(m-1)$, *and* $S^2_{Fp} = \sum_{k=1}^{K}(\hat{\tau}_{\hat{f}_{-k},p}(\mathcal{Z}_k) - \overline{\hat{\tau}_{\hat{f}_{-k},p}(\mathcal{Z}_k)})^2/(K-1)$, *and* $\kappa_{Ft}(p) = \mathbb{E}(\tau_i \mid \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = t)$, *with* $Y_{\hat{f}pi}(t) = \{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) - p\}Y_i(t)$, $\overline{Y_{\hat{f}p}(t)} = \sum_{i=1}^{n} Y_{\hat{f}pi}(t)/n$, *and* $\overline{\hat{\tau}_{\hat{f}_{-k},p}(\mathcal{Z}_k)} = \sum_{k=1}^{K} \hat{\tau}_{\hat{f}_{-k},p}(\mathcal{Z}_k)/K$, *for* $t = 0, 1$.

Proof is given in Appendix A.6. The estimation of the term $\mathbb{E}(\widetilde{S}^2_{\hat{f}t})$ is done similarly as before. For $\kappa_{Ft}(p)$, we replace it with its sample analogue:

$$\widehat{\kappa_{Ft}(p)} \;=\; \frac{1}{K}\sum_{l=1}^{K} \frac{\sum_{i=1}^{m} \mathbf{1}\{\hat{f}_{-k}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{-k})) = t\} T_i^{(k)} Y_i^{(k)}}{\sum_{i=1}^{m} \mathbf{1}\{\hat{f}_{-k}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{-k})) = t\} T_i^{(k)}} - \frac{\sum_{i=1}^{m} \mathbf{1}\{\hat{f}_{-k}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{-k})) = t\}(1 - T_i^{(k)}) Y_i^{(k)}}{\sum_{i=1}^{m} \mathbf{1}\{\hat{f}_{-k}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{-k})) = t\}(1 - T_i^{(k)})}. \tag{16}$$

To estimate the term that appears in the denominator of $\gamma_p(\epsilon)$ as part of the upper bound of bias, we assume that the CATE, i.e., $\mathbb{E}(Y_i(1) - Y_i(0) \mid \hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = c)$, is continuous in $c$, and replace the maximum with a point estimate. We may also utilize an upper bound for CATE, if known, to estimate the bias conservatively. Building on this result, Appendix A.7 shows how to compare two estimated ITRs by estimating the PAPD (Eqn (12)).

Finally, we consider the following cross-validation estimator of the AUPEC for an estimated ITR (Eqn (13)),

$$\widehat{\Gamma}_F \;=\; \frac{1}{K}\sum_{k=1}^{K} \widehat{\Gamma}_{\hat{f}_{-k}}(\mathcal{Z}_k), \tag{17}$$

where $\widehat{\Gamma}_f$ is defined in Eqn (9). This can be seen as an overall statistic that measures the prescriptive performance of the machine learning algorithm $F$ on the dataset under cross-validation. The following theorem derives a bias bound and the exact variance of this cross-validation estimator.

THEOREM 4 (BIAS BOUND AND EXACT VARIANCE OF THE CROSS-VALIDATION AUPEC ESTIMATOR) *Under Assumptions 1, 2, and 3, the bias of the AUPEC estimator defined in Eqn* (17) *can be bounded as follows,*

$$\begin{aligned}\mathbb{E}_{\mathcal{Z}^{tr}}[\mathbb{P}_{\hat{p}_{\hat{f}}}(|\mathbb{E}(\widehat{\Gamma}_F - \Gamma_F \mid \hat{p}_{\hat{f}})| \geq \epsilon)] \;\leq\; & \mathbb{E}\{1 - B(1 - p_{\hat{f}} + \gamma_{p_{\hat{f}}}(\epsilon), m - \lfloor mp_{\hat{f}} \rfloor, \lfloor mp_{\hat{f}} \rfloor + 1) \\ & + B(1 - p_{\hat{f}} - \gamma_{p_{\hat{f}}}(\epsilon), m - \lfloor mp_{\hat{f}} \rfloor, \lfloor mp_{\hat{f}} \rfloor + 1)\},\end{aligned}$$

*where any given constant $\epsilon > 0$, $B(\epsilon, \alpha, \beta)$ is the incomplete beta function (if $\alpha = 0$ and $\beta > 0$, we set $B(\epsilon, \alpha, \beta) := H(\epsilon)$ for all $\epsilon$ where $H(\epsilon)$ is the Heaviside step function), and*

$$\gamma_{p_{\hat{f}}}(\epsilon) \;=\; \frac{\epsilon}{2\mathbb{E}_{\mathcal{Z}^{tr}}\{\max_{c \in [c^* - \epsilon,\, c^* + \epsilon]} \mathbb{E}(\tau_i \mid \hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = c)\}}.$$

*The variance is given by,*

$$\begin{aligned}\mathbb{V}(\widehat{\Gamma}_F) \;=\; & \mathbb{E}\left[-\frac{1}{m}\left\{\sum_{z=1}^{Z} \frac{k(n-z)}{m^2(m-1)}\kappa_{F1}(z/m)\kappa_{F0}(z/m) + \frac{Z(m-Z)^2}{m^2(m-1)}\kappa_{F1}(Z/m)\kappa_{F0}(Z/m)\right\}\right. \\ & - \frac{2}{m^4(m-1)}\sum_{z=1}^{Z-1}\sum_{z'=z+1}^{Z} z(m-z')\kappa_{F1}(z/m)\kappa_{F1}(z'/m) - \frac{Z^2(m-Z)^2}{m^4(m-1)}\kappa_{F1}(Z/m)^2 \\ & \left. - \frac{2(m-Z)^2}{m^4(m-1)}\sum_{z=1}^{Z} k\kappa_{F1}(Z/m)\kappa_{F1}(z/m) + \frac{1}{m^4}\sum_{z=1}^{Z} z(m-z)\kappa_{F1}(z/m)^2\right] \\ & + \mathbb{V}\left(\sum_{z=1}^{Z} \frac{z}{m}\kappa_{F1}(z/m) + \frac{(m-Z)Z}{m}\kappa_{F1}(Z/m)\right) + \frac{\mathbb{E}(S_{\hat{f}1}^{*2})}{m_1} + \frac{\mathbb{E}(S_{\hat{f}0}^{*2})}{m_0} - \frac{K-1}{K}\mathbb{E}(S_F^{*2}),\end{aligned}$$

*where $Z$ is a Binomial random variable with size $m$ and success probability $p_{\hat{f}}$, $S_{\hat{f}t}^{*2} = \sum_{i=1}^{m}(Y_{\hat{f}i}^*(t) - \overline{Y_{\hat{f}}^*(t)})^2/(m-1)$, $S_F^2 = \sum_{k=1}^{K}(\widehat{\Gamma}_{\hat{f}_{-k}}(\mathcal{Z}_k) - \overline{\widehat{\Gamma}_{\hat{f}_{-k}}(\mathcal{Z}_k)})^2/(K-1)$, and $\kappa_{Ft}(z/m) = \mathbb{E}(\tau_i \mid \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_{z/m}(\hat{f}_{\mathcal{Z}^{tr}}) = t)$ with $\overline{Y_{\hat{f}}^*(t)} = \sum_{i=1}^{m} Y_{\hat{f}i}^*(t)/m$, $\overline{\widehat{\Gamma}_{\hat{f}_{-k}}(\mathcal{Z}_k)} = \sum_{k=1}^{K}\widehat{\Gamma}_{\hat{f}_{-k}}(\mathcal{Z}_k)/K$, and $Y_{\hat{f}i}^*(t) = \left[\left\{\sum_{z=1}^{m_f}\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_{z/m}(\hat{f}_{\mathcal{Z}^{tr}})) + (m - m_f)\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_{z/m}(\hat{f}_{\mathcal{Z}^{tr}}))\right\}/m - \frac{1}{2}\right]$ for $t = 0, 1$.*

Proof is similar to that of Theorem 2. The estimation of $\mathbb{E}(S_{\hat{f}1}^{*2})$, $\mathbb{E}(S_{\hat{f}0}^{*2})$, and $\mathbb{E}(S_F^{*2})$ is the same as before, and the $\kappa_{Ft}(p)$ term can be estimated using Eqn (16).

# 5 A Simulation Study

We conduct a simulation study to examine the finite sample performance of our methodology, for both fixed and estimated ITRs. We find that the empirical coverage probability of the confidence interval, based on the proposed variance, approximates its nominal rate even in a small sample. We also find that the bias is minimal even when the proposed estimator is not unbiased and that our variance bounds are tight.

## 5.1 Data Generation Process

Our data generating process (DGP) is based on the one used in the 2017 Atlantic Causal Inference Conference (ACIC) Data Analysis Challenge (Hahn *et al.*, 2018). A total of 8 covariates $\mathbf{X}$ are taken from the Infant Health and Development Program, which originally had 58 covariates and $4,302$ observations. In our simulation, the population distribution of covariates is assumed to equal the empirical distribution of this data set. Therefore, we obtain each simulation sample via boostrap. We vary the sample size: $n = 100, 500, 2000$.

We use the same outcome model as the one used in the competition,

$$\mathbb{E}(Y_i(t) \mid \mathbf{X}_i) \;=\; \mu(\mathbf{X}_i) + \tau(\mathbf{X}_i)t, \tag{18}$$

where $\pi(\mathbf{x}) = 1/[1+\exp\{3(x_1+x_{43}+0.3(x_{10}-1))-1\}]$, $\mu(\mathbf{x}) = -\sin(\Phi(\pi(\mathbf{x})))+x_{43}$, and $\tau(\mathbf{x}) = \xi(x_3x_{24}+(x_{14}-1)-(x_{15}-1))$ with $\Phi(\cdot)$ representing the standard Normal CDF and $x_j$ indicating a specific covariate in the data set. One important difference is that we assume a complete randomized experiment whereas the original DGP generated the treatment using a function of covariates to emulate an observational study. As in the competition, we focus on two scenarios regarding the treatment effect size by setting $\xi$ equal to $2$ ("high") and $1/3$ ("low"). Although the original DGP included four different error distributions, we use the i.i.d. error, $\sigma(\mathbf{X}_i)\epsilon_i$ where $\sigma(\mathbf{x}) = 0.25\sqrt{\mathbb{V}(\mu(\mathbf{x}) + \pi(\mathbf{x})\tau(\mathbf{x}))}$ and $\epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$.

For fixed ITRs, we can directly compute the true values of our causal quantities of interest using the outcome model specified in Eqn (18) and evaluate each quantity based on the entire original data set. This computation is valid because we assume the population distribution of covariates is equal to the empirical distribution of the original data set. For the estimated ITR case, however, we do not have an analytical expression for the true value of a causal quantity of interest. Therefore, we obtain an approximate true value via Monte Carlo simulation. We generate 10,000 independent datasets based on the same DGP, and train the specified algorithm $F$ on each of the datasets using 5-fold cross-validation (i.e., $K = 5$). Then, we use the sample mean of our estimated causal quantity across 10,000 data sets as our approximate truth.

We evaluate Bayesian Additive Regression Trees (BART) (Chipman *et al.*, 2010; Hahn *et al.*, 2020), which had the best overall performance in the original competition. We compare this model with two other popular methods: Causal Forest (Athey *et al.*, 2019) as well as the LASSO, which includes all main effects and two-way interaction effects between the treatment and all covariates (Tibshirani, 1996). All three models are trained on the original data from the 2017 ACIC Data Challenge. The number of trees was tuned through the 5-fold cross validation for BART and Causal Forest. The regularization parameter was tuned similarly for LASSO. All models were cross-validated on the PAPE. For implementation, we use R 3.4.2 with bartMachine (version 1.4.2) for BART, grf (version 0.10.2) for Causal Forest, and glmnet (version 2.0.13) for LASSO. Once the models are trained, an ITR is derived based on the magnitude of the estimated CATE $\hat{\tau}(\mathbf{x})$, i.e., $f(\mathbf{X}_i) = \mathbf{1}\{\hat{\tau}(\mathbf{X}_i) > 0\}$.

## 5.2 Results

We first present the results for fixed ITRs followed by those for estimated ITRs. Table 1 presents the bias and standard deviation of each estimator for fixed ITRs as well as the coverage probability of its 95% confidence intervals based on 1,000 Monte Carlo trials. The results are shown separately for the high and low treatment effects scenarios. We estimate the PAPE $\tau_f$ for BART without a budget constraint as well as the PAPE with a budget constraint of 20% as the maximal proportion of treated units, $\tau_f(c_{0.2})$. In addition, we estimate the AUPEC $\Gamma_f$ and compute the difference in the PAPE or PAPD between BART and Causal Forest ($\Delta(f, g)$), and between BART and LASSO ($\Delta(f, h)$).

Under both scenarios and across sample sizes, the bias of our estimator is small. More-

| Estimator | truth | $n = 100$ coverage | bias | s.d. | $n = 500$ coverage | bias | s.d. | $n = 2000$ coverage | bias | s.d. |
|---|---|---|---|---|---|---|---|---|---|---|
| **Low treatment effect** | | | | | | | | | | |
| $\hat{\tau}_f$ | 0.066 | 94.3% | 0.005 | 0.124 | 96.2% | 0.001 | 0.053 | 95.1% | 0.001 | 0.026 |
| $\hat{\tau}_f(c_{0.2})$ | 0.051 | 93.2 | −0.002 | 0.109 | 94.4 | 0.001 | 0.046 | 95.2 | 0.002 | 0.021 |
| $\widehat{\Gamma}_f$ | 0.053 | 95.3 | 0.001 | 0.106 | 95.1 | 0.001 | 0.045 | 94.8 | −0.001 | 0.024 |
| $\widehat{\Delta}_{0.2}(f, g)$ | −0.022 | 94.0 | 0.006 | 0.122 | 95.4 | 0.002 | 0.051 | 96.0 | 0.000 | 0.026 |
| $\widehat{\Delta}_{0.2}(f, h)$ | −0.014 | 93.9 | −0.001 | 0.131 | 94.9 | −0.000 | 0.060 | 95.3 | −0.000 | 0.030 |
| **High treatment effect** | | | | | | | | | | |
| $\hat{\tau}_f$ | 0.430 | 94.7% | −0.000 | 0.163 | 95.7% | 0.000 | 0.064 | 94.4% | −0.000 | 0.031 |
| $\hat{\tau}_f(c_{0.2})$ | 0.356 | 94.7 | 0.004 | 0.159 | 95.7 | 0.002 | 0.072 | 95.8 | 0.000 | 0.035 |
| $\widehat{\Gamma}_f$ | 0.363 | 94.3 | −0.005 | 0.130 | 94.9 | 0.003 | 0.058 | 95.7 | 0.000 | 0.029 |
| $\widehat{\Delta}_{0.2}(f, g)$ | −0.000 | 96.9 | 0.008 | 0.151 | 97.9 | −0.002 | 0.073 | 98.0 | −0.000 | 0.026 |
| $\widehat{\Delta}_{0.2}(f, h)$ | 0.000 | 94.7 | −0.004 | 0.140 | 97.7 | −0.001 | 0.065 | 96.6 | 0.000 | 0.033 |

Table 1: The Results of the Simulation Study for Fixed Individualized Treatment Rules (ITRs). The table presents the bias and standard deviation of each estimator as well as the coverage of its 95% confidence intervals under the "Low treatment effect" and "High treatment effect" scenarios. The first three estimators shown here are for BART $f$: Population Average Prescription effect (PAPE; $\hat{\tau}_f$), PAPE with a budget constraint of 20% treatment proportion ($\hat{\tau}_f(c_{0.2})$), Area Under the Prescriptive Effect Curve (AUPEC; $\widehat{\Gamma}_f$). We also present the results for the difference in the PAPE between BART and Causal Forest $g$ ($\widehat{\Delta}_{0.2}(f, g)$) and between BART and LASSO $h$ ($\widehat{\Delta}_{0.2}(f, g)$) under the budget constraint.

| Estimator | $n = 100$ truth | coverage | bias | s.d. | $n = 500$ truth | coverage | bias | s.d. | $n = 2000$ truth | coverage | bias | s.d. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Low Effect** | | | | | | | | | | | | |
| $\hat{\lambda}_F$ | 0.073 | 96.4% | 0.001 | 0.216 | 0.095 | 96.7% | 0.002 | 0.100 | 0.112 | 97.2% | 0.002 | 0.046 |
| $\hat{\tau}_F$ | 0.021 | 94.6 | −0.002 | 0.130 | 0.030 | 95.5 | −0.002 | 0.052 | 0.032 | 94.4 | −0.000 | 0.027 |
| $\hat{\tau}_F(c_{0.2})$ | 0.023 | 95.4 | −0.003 | 0.120 | 0.034 | 95.4 | −0.002 | 0.057 | 0.043 | 96.8 | 0.001 | 0.029 |
| $\widehat{\Gamma}_F$ | 0.009 | 98.2 | 0.002 | 0.117 | 0.029 | 96.8 | −0.001 | 0.048 | 0.039 | 95.9 | 0.001 | 0.001 |
| **High Effect** | | | | | | | | | | | | |
| $\hat{\lambda}_H$ | 0.867 | 96.9% | −0.007 | 0.261 | 0.875 | 96.5% | −0.003 | 0.125 | 0.875 | 97.3% | 0.001 | 0.062 |
| $\hat{\tau}_F$ | 0.338 | 93.6 | −0.000 | 0.171 | 0.358 | 93.0 | 0.000 | 0.093 | 0.391 | 95.3 | 0.001 | 0.041 |
| $\hat{\tau}_F(c_{0.2})$ | 0.341 | 94.8 | −0.002 | 0.170 | 0.356 | 96.2 | −0.005 | 0.075 | 0.356 | 95.8 | 0.001 | 0.037 |
| $\widehat{\Gamma}_F$ | 0.344 | 98.5 | 0.001 | 0.126 | 0.362 | 98.9 | 0.005 | 0.053 | 0.363 | 99.0 | 0.001 | 0.026 |

Table 2: The Results of the Simulation Study for Cross-validated ITR. The table presents the true value (truth) of each quantity along with the bias and standard deviation of each estimator as well as the coverage of its 95% confidence intervals under the "Low treatment effect" and "High treatment effect" scenarios. All of the results shown here are for LASSO.

over, the coverage rate of 95% confidence intervals is close to their nominal rate even when the sample size is small. Although we can only bound the variance when estimating the PAPD between two ITRs (i.e., $\Delta_{0.2}(f, g)$ and $\Delta_{0.2}(f, h)$), the coverage stay close to 95%, implying that the bound for covariance has little effect on the variance estimation.

For estimated ITRs, Table 2 presents the results of LASSO under cross-validation. For

BART and Causal Forest, obtaining an accurate Monte Carlo estimate of the true causal parameter values under cross-validation takes a prohibitively large amount of time. While the out-of-bag estimates of such true values can be computed, they have been shown to create bias under certain scenarios (Janitza and Hornung, 2018). These true values are generally greater for a larger sample size because LASSO performs better with more data.

The proposed cross-validated estimators are approximately unbiased even for $n = 100$. The coverage is generally around or above the nominal $95\%$ value, reflecting the conservative estimate of the variance. For the PAPE without and with budget constraint, i.e., $\hat{\tau}_F$ and $\hat{\tau}_F(c_{0.2})$, the coverage is close to the nominal value. This indicates that the bias of the proposed conservative variance estimator is relatively small even though the number of folds for cross-validation is only $K = 5$. The performance of the proposed methodology is good even when the sample size is as small as $n = 100$. When the sample size is $n = 500$, the standard deviation of the cross-validated estimator is roughly half of the corresponding $n = 100$ fixed ITR estimator (this is a good comparison because each fold has 100 observations). This confirms the theoretical efficiency gain that results from cross-validation.

# 6 An Empirical Application

We apply the proposed methodology to the data from the Tennessee's Student/Teacher Achievement Ratio (STAR) project, which was a longitudinal study experimentally evaluating the impacts of class size in early education on various outcomes (Mosteller, 1995). Another application based on a canvassing experiment is shown in Appendix A.8.

## 6.1 Data and Setup

The STAR project randomly assigned over 7,000 students across 79 schools to three different groups: small class,regular class,and regular class with a full-time teacher's aid. The experiment began when students entered kindergarden and continued through third grade. To create a binary treatment, we focus on the first two groups: small class and regular class without an aid. The treatment effect heterogeneity is important because reducing class size is costly, requiring additional teachers and classrooms. Policy makers who face a budget constraint may be interested in finding out which groups of students benefit most from a small class size so that the priority can be given to those students.

We follow the analysis strategies of the previous studies (e.g., Ding and Lehrer, 2011; McKee *et al.*, 2015) that estimated the heterogeneous effects of small classes on educational attainment. These authors adjust for school-level and student-level characteristics, but do not consider within-classroom interactions. Unfortunately, addressing this limitation is beyond the scope of this paper. We use a total of 10 pre-treatment covariates $\mathbf{X}_i$ that include four demographic characteristics of students (gender, race, birth month, birth year) and six school characteristics (urban/rural, enrollment size, grade range, number of students on free lunch, number of students on school buses, and percentage of white students). Our treatment variable is the class size to which they were assigned at kindergarten: small class $T_i = 1$ and regular class without an aid $T_i = 0$. For the outcome variables $Y_i$, we use three standardized test scores measured at third grade: math, reading, and writing SAT scores.

The resulting data set has a total of 1,911 observations. The estimated average treatment effects (ATE) based on the entire data set are $6.78$ (s.e. = $1.71$), $5.78$ (s.e. = $1.80$), and $3.65$ (s.e. = $1.63$), for the reading, math, and writing scores, respectively. For the fixed test data, the estimated ATEs are similar; $5.10$ (s.e. = $3.07$), $2.78$ (s.e.= $3.15$), and $1.48$ (s.e. = $2.96$).

We evaluate the performance of ITRs using two settings considered above. First, we randomly split the data into the training data (70%) and test data (30%). We estimate an ITR from the training data and then evaluate it as a fixed ITR using the test data. This follows the setup considered in Sections 2 and 3. Second, we consider the evaluation of estimated ITRs based on the same experimental data. We utilize Algorithm 1 with 5-fold cross-validation (i.e., $K = 5$). For both settings, we use the same three machine learning algorithms. For Causal Forest, we set `tune.parameters = TRUE`. For BART, tuning

| | BART | | | Causal Forest | | | LASSO | | |
|---|---|---|---|---|---|---|---|---|---|
| | est. | s.e. | treated | est. | s.e. | treated | est. | s.e. | treated |
| **Fixed ITR** | | | | | | | | | |
| *No budget constraint* | | | | | | | | | |
| Reading | 0 | 0 | 100% | −0.38 | 1.14 | 84.3% | −0.41 | 1.10 | 84.4% |
| Math | 0.52 | 1.09 | 86.7 | 0.09 | 1.18 | 80.3 | 1.73 | 1.25 | 78.7 |
| Writing | −0.32 | 0.72 | 92.7 | −0.70 | 1.18 | 78.0 | −0.30 | 1.26 | 80.0 |
| *Budget constraint* | | | | | | | | | |
| Reading | −0.89 | 1.30 | 20 | 0.66 | 1.23 | 20 | −1.17 | 1.18 | 20 |
| Math | 0.70 | 1.25 | 20 | 2.57 | 1.29 | 20 | 1.25 | 1.32 | 20 |
| Writing | 2.60 | 1.17 | 20 | 2.98 | 1.18 | 20 | 0.28 | 1.19 | 20 |
| **Estimated ITR** | | | | | | | | | |
| *No budget constraint* | | | | | | | | | |
| Reading | 0.19 | 0.37 | 99.3% | 0.31 | 0.77 | 86.6% | 0.32 | 0.53 | 87.6% |
| Math | 0.92 | 0.75 | 84.7 | 2.29 | 0.80 | 79.1 | 1.52 | 1.60 | 75.2 |
| Writing | 1.12 | 0.86 | 88.0 | 1.43 | 0.71 | 67.4 | 0.05 | 1.37 | 74.8 |
| *Budget constraint* | | | | | | | | | |
| Reading | 1.55 | 1.05 | 20 | 0.40 | 0.69 | 20 | −0.15 | 1.41 | 20 |
| Math | 2.28 | 1.15 | 20 | 1.84 | 0.73 | 20 | 1.50 | 1.48 | 20 |
| Writing | 2.31 | 0.66 | 20 | 1.90 | 0.64 | 20 | −0.47 | 1.34 | 20 |

Table 3: The Estimated Population Average Prescription Effect (PAPE) for BART, Causal Forest, and LASSO with and without a Budget Constraint. We estimate the PAPE for fixed and estimated individualized treatment rules (ITRs). The fixed ITRs are based on the training (70%) and test data (30%), whereas the estimated ITRs are based on 5 fold cross-validation. In addition, the average treatment effect estimates using the entire dataset. For each of the three outcomes, the point estimate, the standard error, and the average proportion treated are shown. The budget constraint considered here implies that the maximum proportion treated is 20%.

was done on the number of trees. For LASSO, we tuned the regularization parameter while including all interaction terms between covariates and the treatment variable. All tuning was done through the 5-fold cross validation procedure on the training set using the PAPE as the evaluation metric. We then create an ITR as $\mathbf{1}\{\hat{\tau}(\mathbf{x}) > 0\}$ where $\hat{\tau}(\mathbf{x})$ is the estimated CATE obtained from each fitted model. As mentioned in Section 3, we center the outcome variable $Y$ in evaluating the metrics to minimize the variance of the estimators.

## 6.2 Results

The upper panel of Table 3 presents the estimated PAPEs, their standard errors, and the proportion treated for fixed ITRs. We find that without a budget constraint, none of the

| | Causal Forest | | | | BART | |
|---|---|---|---|---|---|---|
| | vs. **BART** | | vs. **LASSO** | | vs. **LASSO** | |
| | est. | 95% CI | est. | 95% CI | est. | 95% CI |
| **Fixed ITR** | | | | | | |
| Math | 1.55 | [−0.35, 3.45] | 1.83 | [−0.50, 4.16] | 0.28 | [−2.39, 2.95] |
| Reading | 1.86 | [−0.79, 4.51] | 1.31 | [−1.49, 4.11] | −0.55 | [−4.02, 2.92] |
| Writing | 0.38 | [−1.66, 2.42] | 2.69 | [−0.27, 5.65] | 2.32 | [−0.53, 5.15] |
| **Estimated ITR** | | | | | | |
| Reading | −1.15 | [−3.99, 1.69] | 0.55 | [−1.05, 2.15] | 1.70 | [−0.90, 4.30] |
| Math | −0.43 | [−2.57, 3.43] | 0.34 | [−1.32, 2.00] | 0.77 | [−1.99, 3.53] |
| Writing | −0.41 | [−1.63, 0.80] | 2.37 | [0.76, 3.98] | 2.79 | [1.32, 4.26] |

Table 4: The Estimated Population Average Prescriptive Effect Difference (PAPD) between Causal Forest, BART, and LASSO under a Budget Constraint of 20%. The point estimates and 95% confidence intervals are shown.

machine learning algorithms significantly improves upon the random treatment rule. With a budget constraint of 20%, however, the ITRs based on Causal Forest appear to improve upon the random treatment rule at least for the math and writing scores. In contrast, the ITR based on BART only performs well for the writing score whereas the performance of LASSO is not distinguishable from that of random treatment rule across all scores. The results are largely similar for the estimated ITRs, as shown in the lower panel of Table 3. The only difference is that BART performs slightly better while the performance of Causal Forest is slightly better for the fixed ITRs and worse for the estimated ITRs.

In the case of BART and Causal Forest, the standard errors for estimated ITRs are generally smaller than those for fixed ITRs, reflecting the efficiency gain due to cross-validation. For LASSO, however, the standard errors for fixed ITRs are often smaller than those for estimated ITRs. This is due to the fact that the ITRs estimated by LASSO are significantly variable across different folds under cross-validation. This variability results in the poor performance of LASSO in this application. In contrast, Causal Forest is most stable, generally leading to the best performance and the smallest standard errors. Lastly, we note that when compared to the estimated ATEs, some estimated PAPEs are of substantial size.

Table 4 directly compares these three ITRs based on Causal Forest, BART, and LASSO by estimating the Population Average Prescriptive Effect Difference (PAPD) under the

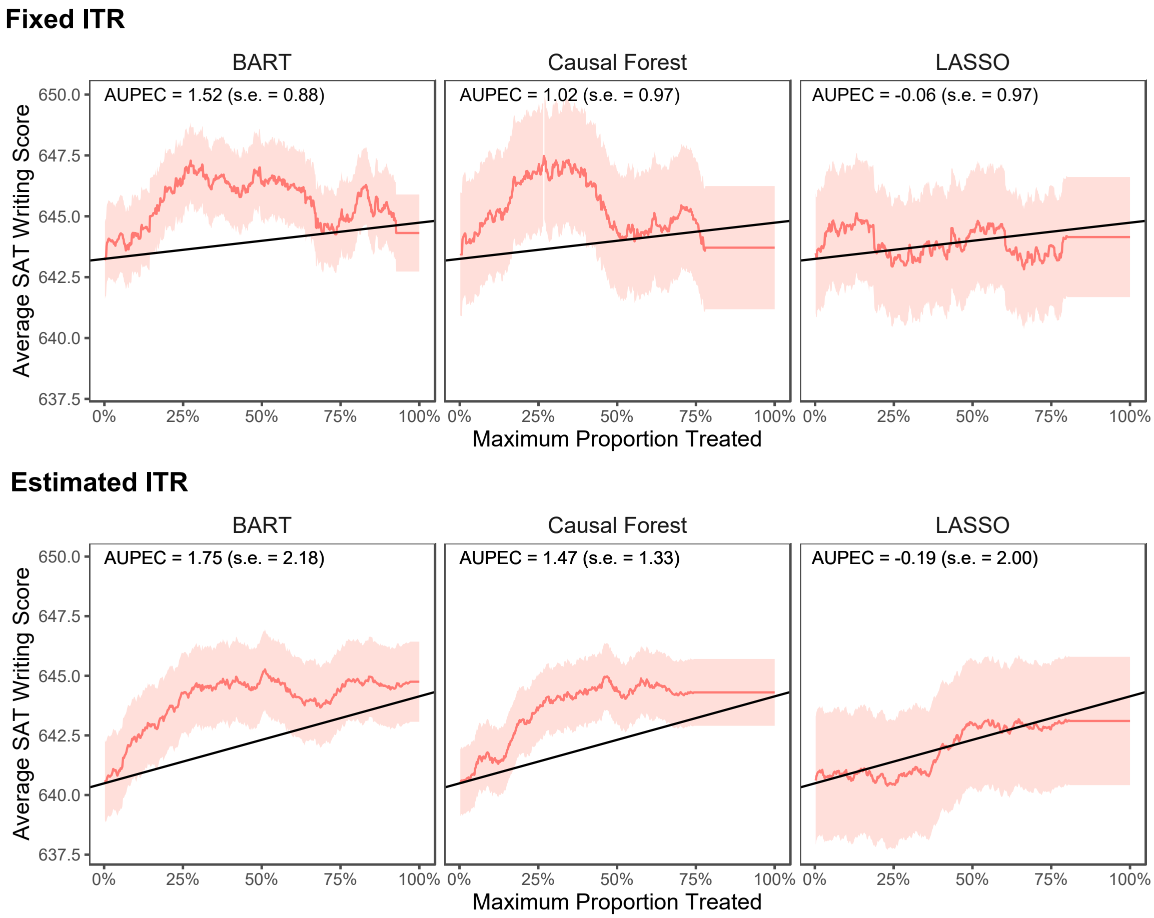


Figure 3: Estimated Area Under the Prescriptive Effect Curve (AUPEC). The results are presented for the fixed (upper panel) and estimated (bottom panel) individualized treatment rule (ITR) settings. A solid red line in each plot represents the Population Average Prescriptive Effect for SAT writing scores across a range of budget constraint (horizontal axis) with the pointwise 95% confidence intervals. The area between this line and the black line representing random treatment is the AUPEC. The results are presented for the individualized treatment rules based on BART (left column), Causal Forest (middle column), and LASSO (right column), whereas each row presents the results for a different outcome.

same budget constraint (i.e., 20%) as above. Causal Forest outperforms BART and LASSO in essentially all cases though the difference is not statistically significant. Under the cross-validation setting, Causal Forest and BART are statistically significantly more effective than LASSO in identifying students with grater treatment effects on their writing scores.

Finally, Figure 3 presents the estimated PAPE for the writing score across a range of budget constraints with the pointwise confidence intervals based on the variance formulas, for the fixed and estimated ITR cases. The difference between the solid red and black lines is the estimated PAPE, while the area between the red line and the black line corresponds to the Area Under the Prescriptive Effect Curve (AUPEC). In each plot, the horizontal axis represents the budget constraint as the maximum proportion treated. In both the fixed and estimated ITR settings, BART and Causal Forest identify students who benefit positively from small class sizes when the maximum proportion treated is relatively small. In contrast, LASSO has a difficult time in finding these individuals. Additionally, we find that the standard error of the estimated AUPEC is greater for estimated ITRs than for fixed ITRs even though in the case of BART and Causal Forest, the opposite pattern is found for the PAPE with a given budget constraint. This finding is due to the high variance of the estimated AUPEC for different folds of cross-validation.

As the budget constraint is relaxed, the ITRs based on BART and Causal Forest yields the population average value similar to the one under the random treatment rule. This results in the inverted V-shape AUPEC curves observed in the left and middle plots of the figure. These inverted V-shape curves illustrate two separate phenomena. First, the students with the highest predicted CATE under BART and Causal Forests do have a higher treatment effect than the average, yielding an uplift in the PAPE curve compared to the random treatment rule. This shows that BART and Causal Forests are able to capture some causal heterogeneity that exists in the STAR data. Indeed, both BART and the Causal Forest estimated the CATE to be higher for non-white students and those who attend schools with a high percentage of students receiving free lunch. According to the variable importance statistic, these two covariates play an essential role in explaining causal heterogeneity (see e.g., Finn and Achilles, 1990; Jackson and Page, 2013; Nye *et al.*, 2000, similar findings).

However, as we further relax the budget constraint, the methods start treating students who are predicted to have a smaller (yet still positive) value of the CATE. These students tend to benefit less than the ATE, resulting in a smaller value of the PAPE as the budget increases. Eventually, both BART and Causal Forest start "over-treating" students who are estimated to have a small positive CATE but actually do not benefit from a small class. This results in the overall insignificant PAPE when no budget constraint is imposed.

## 7 Concluding Remarks

As the application of individualized treatment rules (ITRs) becomes more widespread in a variety of fields, a rigorous perfomance evaluation of ITRs becomes essential before policy makers deploy them in a target population. We believe that the inferential approach proposed in this paper provides a robust and widely applicable tool to policy makers. The proposed methodology also opens up opportunities to utilize the existing randomized controlled trial data for the efficient evaluation of ITRs as done in our empirical application. In addition, although we do not focus on the estimation of ITRs in this paper, the proposed evaluation metrics can be used to tune hyper-parameters when cross validating machine learning algorithms. In future research, we plan to consider the extensions of the proposed methodology to other settings, including non-binary treatments, dynamic treatments, and treatment allocations in the presence of interference between units.

## References


Andrews, I., Kitagawa, T., and McCloskey, A. (2020). Inference on winners. Tech. rep., Department of Economics, Harvard University.

Ascarza, E. (2018). Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research* **55**, 1, 80–98.

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences* **113**, 27, 7353–7360.

Athey, S., Tibshirani, J., Wager, S., *et al.* (2019). Generalized random forests. *Annals of Statistics* **47**, 2, 1148–1178.

Athey, S. and Wager, S. (2018). Efficient policy learning. *arXiv:1702.02896* .

Bohrnstedt, G. W. and Goldberger, A. S. (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association* **64**, 328, 1439–1442.

Broockman, D. and Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science* **352**, 6282, 220–224.

Chakraborty, B., Laber, E., and Zhao, Y.-Q. (2014). Inference about the expected performance of a data-driven dynamic treatment regime. *Clinical Trials* **11**, 4, 408–417.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I. (2019). Generic machine learning inference on heterogeneous treatment effects in randomized experiments. Tech. rep., arXiv:1712.04802.

Chipman, H. A., George, E. I., McCulloch, R. E., *et al.* (2010). Bart: Bayesian additive regression trees. *The Annals of Applied Statistics* **4**, 1, 266–298.

Diemert, E., Betlei, A., Renaudin, C., and Amini, M.-R. (2018). A large scale benchmark for uplift modeling. In *KDD*.

Ding, W. and Lehrer, S. F. (2011). Experimental estimates of the impacts of class size on test scores: Robustness and heterogeneity. *Education Economics* **19**, 3, 229–252.

Dudík, M., Langford, J., and Li, L. (2011). Doubly robust policy evaluation and learning. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, 1097–1104, USA. Omnipress.

Fifield, B. (2018). *Empirical and Computational Methods for Electoral Politics*. Ph.D. thesis, Princeton University.

Finn, J. D. and Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal* **27**, 3, 557–577.

Fu, H., Zhou, J., and Faries, D. E. (2016). Estimating optimal treatment regimes via subgroup identification in randomized control trials and observational studies. *Statistics in medicine* **35**, 19, 3285–3302.

Grimmer, J., Messing, S., and Westwood, S. J. (2017). Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods. *Political Analysis* **25**, 413–434.

Gutierrez, P. and Gérardy, J.-Y. (2016). Causal inference and uplift modelling: A review of the literature. In *Proceedings of Machine Learning Research*, vol. 67: International Conference on Predictive Applications and APIs, 1–13.

Hahn, P. R., Dorie, V., and Murray, J. S. (2018). Atlantic causal inference conference (acic) data analysis challenge 2017. Tech. rep., School of Mathematical and Statistical Sciences. `https://math.la.asu.edu/~prhahn/debrief.pdf`.

Hahn, P. R., Murray, J. S., Carvalho, C. M., *et al.* (2020). Bayesian regression tree models for causal inference: regularization, confounding, and heterogeneous effects. *Bayesian Analysis* **15**, 3, 965–1056.

Hamburg, M. A. and Collins, F. S. (2010). The path to personalized medicine. *New England Journal of Medicine* **363**, 301–304.

Imai, K. and Ratkovic, M. (2013). Estimating treatment effect heterogeneity in randomized program evaluation. *Annals of Applied Statistics* **7**, 1, 443–470.

Imai, K. and Strauss, A. (2011). Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign. *Political Analysis* **19**, 1, 1–19.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Jackson, E. and Page, M. E. (2013). Estimating the distributional effects of education reforms: A look at project star. *Economics of Education Review* **32**, 92–103.

Janitza, S. and Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PloS one* **13**, 8, e0201904.

Jiang, N. and Li, L. (2016). Doubly robust off-policy value evaluation for reinforcement learning. In *Proceedings of The 33rd International Conference on Machine Learning*, vol. 48 of *Proceedings of Machine Learning Research*, 652–661.

Kallus, N. (2018). Balanced policy evaluation and learning. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, eds., *Advances in Neural Information Processing Systems 31*, 8895–8906. Curran Associates, Inc.

Kamath, P. S., Wiesner, R. H., Malinchoc, M., Kremers, W., Therneau, T. M., Kosberg, C. L., D'Amico, G., Dickson, E. R., and Kim, W. R. (2001). A model to predict survival in patients with end-stage liver disease. *Hepatology* **33**, 2, 464–470.

Kitagawa, T. and Tetenov, A. (2018). Who should be treated?: Empirical welfare maximization methods for treatment choice. *Econometrica* **86**, 2, 591–616.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., and Yu, B. (2018). Meta-learners for estimating heterogeneous treatment effects using machine learning. Tech. rep., arXiv:1706.03461.

Luedtke, A. R. and van der Laan, M. J. (2016a). Optimal individualized treatments in resource-limited settings. *International Journal of Biostatistics* **12**, 1, 283–303.

Luedtke, A. R. and van der Laan, M. J. (2016b). Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *Annals of Statistics* **44**, 2, 713–742.

Manski, C. F. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica* **72**, 4, 1221–1246.

McKee, G., Sims, K. R. E., and Rivkin, S. G. (2015). Disruption, learning, and the heterogeneous benefits of smaller classes. *Empirical Economics* **48**, 3, 1267–1286.

Mosteller, F. (1995). The tennessee study of class size in the early school grades. *The Future of Children* **5**, 2, 113–127.

Nadeau, C. and Bengio, Y. (2000). Inference for the generalization error. In *Advances in neural information processing systems*, 307–313.

Neyman, J. (1923). On the application of probability theory to agricultural experiments: Essay on principles, section 9. (translated in 1990). *Statistical Science* **5**, 465–480.

Nie, X. and Wager, S. (2017). Quasi-oracle estimation of heterogeneous treatment effects. *arXiv preprint arXiv:1712.04912* .

Nye, B. A., Hedges, L. V., and Konstantopoulos, S. (2000). Do the disadvantaged benefit more from small classes? evidence from the tennessee class size experiment. *American journal of education* **109**, 1, 1–26.

Qian, M. and Murphy, S. A. (2011). Performance gurantees for individualized treatment rules. *Annals of Statistics* **39**, 2, 1180–1210.

Radcliffe, N. J. (2007). Using control groups to target on predicted lift: Building and assessing uplift models. *Direct Marketing Analytics Journal* **1**, 3, 14–21.

Rai, Y. (2018). Statistical inference for treatment assignment policies. Tech. rep., Department of Economics, University of Wisconsin.

Rzepakowski, P. and Jaroszewicz, S. (2012). Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems* **32**, 2, 303–327.

Stuard, A. and Ord, J. K. (1994). *Kendall's Advanced Theory of Statistics. Volume 1. Distribution Theory*. Halsted Press, New York, NY.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**, 1, 267–288.

van der Laan, M. J. and Rose, S. (2011). *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association* **113**, 523, 1228–1242.

Zhang, B., Tsiatis, A. A., Davidian, M., and Laber, E. (2012). Estimating optimal treatment regimes from a classification perspective. *Stat* **1**, 1, 103–114.

Zhao, Y., Zeng, D., Rush, J. A., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107**, 499, 1106–1118.

Zhou, X., Mayer-Hamblett, N., Khan, U., and Kosorok, M. R. (2017). Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association* **112**, 517, 169–187.

Zhou, Z., Athey, S., and Wager, S. (2018). Offline multi-action policy learning: Generalization and optimization. Tech. rep., arXiv: 1810.04778.

# A Supplementary Appendix for "Experimental Evaluation of Individualized Treatment Rules"

## A.1 Estimation and Inference for Fixed ITRs with No Budget Constraint

### A.1.1 The Population Average Value

For a fixed ITR with no budget constraint, the following unbiased estimator of the population average value (equation (1)), based on the experimental data $\mathcal{Z}$, is used in the literature (e.g., Qian and Murphy, 2011),

$$\hat{\lambda}_f(\mathcal{Z}) \;=\; \frac{1}{n_1}\sum_{i=1}^{n} Y_i T_i f(\mathbf{X}_i) + \frac{1}{n_0}\sum_{i=1}^{n} Y_i(1-T_i)(1-f(\mathbf{X}_i)). \tag{A1}$$

Under Neyman's repeated sampling framework, it is straightforward to derive the unbiasedness and variance of this estimator where the uncertainty is based solely on the random sampling of units and the randomization of treatment alone. The results are summarized as the following theorem.

THEOREM A1 (UNBIASEDNESS AND VARIANCE OF THE POPULATION AVERAGE VALUE ESTIMATOR) *Under Assumptions 1,2, and 3, the expectation and variance of the population average value estimator defined in equation* (A1) *are given by,*

$$\mathbb{E}\{\hat{\lambda}_f(\mathcal{Z})\} - \lambda_f \;=\; 0, \quad \mathbb{V}\{\hat{\lambda}_f(\mathcal{Z})\} \;=\; \frac{\mathbb{E}(S_{f1}^2)}{n_1} + \frac{\mathbb{E}(S_{f0}^2)}{n_0}$$

*where* $S_{ft}^2 = \sum_{i=1}^{n}(Y_{fi}(t) - \overline{Y_f(t)})^2/(n-1)$ *with* $Y_{fi}(t) = \mathbf{1}\{f(\mathbf{X}_i) = 1\}Y_i(t)$ *and* $\overline{Y_f(t)} = \sum_{i=1}^{n} Y_{fi}(t)/n$ *for* $t = \{0,1\}$.

Proof is straightforward and hence omitted.

### A.1.2 The Population Average Prescriptive Effect (PAPE)

To estimate the PAPE with no budget constraint (equation (2)), we propose the following estimator,

$$\begin{aligned}\hat{\tau}_f(\mathcal{Z}) \;=\; \frac{n}{n-1}\Bigg[&\frac{1}{n_1}\sum_{i=1}^{n} Y_i T_i f(\mathbf{X}_i) + \frac{1}{n_0}\sum_{i=1}^{n} Y_i(1-T_i)(1-f(\mathbf{X}_i)) \\ &-\frac{\hat{p}_f}{n_1}\sum_{i=1}^{n} Y_i T_i - \frac{1-\hat{p}_f}{n_0}\sum_{i=1}^{n} Y_i(1-T_i)\Bigg]\end{aligned} \tag{A2}$$

where $\hat{p}_f = \sum_{i=1}^{n} f(\mathbf{X}_i)/n$ is a sample estimate of $p_f$, and the term $n/(n-1)$ is due to the finite sample degree-of-freedom correction resulting from the need to estimate $p_f$. The following theorem proves the unbiasedness of this estimator and derives its exact variance.

THEOREM A2 (UNBIASEDNESS AND EXACT VARIANCE OF THE PAPE ESTIMATOR) *Under Assumptions 1, 2, and 3, the expectation and variance of the PAPE estimator defined equation* (A2) *are given by,*

$$\begin{aligned}\mathbb{E}\{\hat{\tau}_f(\mathcal{Z})\} &= \tau_f \\ \mathbb{V}\{\hat{\tau}_f(\mathcal{Z})\} &= \frac{n^2}{(n-1)^2}\left[\frac{\mathbb{E}(\widetilde{S}_{f1}^2)}{n_1}+\frac{\mathbb{E}(\widetilde{S}_{f0}^2)}{n_0}+\frac{1}{n^2}\left\{\tau_f^2-np_f(1-p_f)\tau^2+2(n-1)(2p_f-1)\tau_f\tau\right\}\right]\end{aligned}$$

*where* $\widetilde{S}_{ft}^2=\sum_{i=1}^n(\widetilde{Y}_{fi}(t)-\overline{\widetilde{Y}_f(t)})^2/(n-1)$ *with* $\widetilde{Y}_{fi}(t)=(f(\mathbf{X}_i)-\hat{p}_f)Y_i(t)$, *and* $\overline{\widetilde{Y}_f(t)}=\sum_{i=1}^n\widetilde{Y}_{fi}(t)/n$ *for* $t=\{0,1\}$.

Note that $\mathbb{E}(\widetilde{S}_{ft}^2)$ does not equal $\mathbb{V}(\widetilde{Y}_{fi}(t))$ because the proportion of treated units $p_f$ is estimated. The additional term in the variance accounts for this estimation uncertainty of $p_f$. The variance of the proposed estimator can be consistently estimated by replacing the unknown terms, i.e., $p_f$, $\tau_f$, $\tau$, $\mathbb{E}(\widetilde{S}_{ft}^2)$, with their unbiased estimates, i.e., $\hat{p}_f$, $\hat{\tau}_f$,

$$\hat{\tau} = \frac{1}{n_1}\sum_{i=1}^n T_iY_i-\frac{1}{n_0}\sum_{i=1}^n(1-T_i)Y_i, \quad \text{and} \quad \widehat{\mathbb{E}(\widetilde{S}_{ft}^2)} = \frac{1}{n_t-1}\sum_{i=1}^n \mathbf{1}\{T_i=t\}(\widetilde{Y}_{fi}-\overline{\widetilde{Y}_{ft}})^2,$$

where $\widetilde{Y}_{fi}=(f(\mathbf{X}_i)-\hat{p}_f)Y_i$ and $\overline{\widetilde{Y}_{ft}}=\sum_{i=1}^n\mathbf{1}\{T_i=t\}\widetilde{Y}_{fi}/n_t$.

To prove Theorem A2, we first consider the sample average prescription effect (SAPE),

$$\tau_f^s = \frac{1}{n}\sum_{i=1}^n\left\{Y_i(f(\mathbf{X}_i))-\hat{p}_fY_i(1)-(1-\hat{p}_f)Y_i(0)\right\}. \tag{A3}$$

and its unbiased estimator,

$$\hat{\tau}_f^s = \frac{1}{n_1}\sum_{i=1}^n Y_iT_if(\mathbf{X}_i)+\frac{1}{n_0}\sum_{i=1}^n Y_i(1-T_i)(1-f(\mathbf{X}_i))-\frac{\hat{p}_f}{n_1}\sum_{i=1}^n Y_iT_i-\frac{1-\hat{p}_f}{n_0}\sum_{i=1}^n Y_i(1-T_i) \tag{A4}$$

This estimator differs from the estimator of the PAPE by a small factor, i.e., $\hat{\tau}_f^s=(n-1)/n\hat{\tau}_f$. The following lemma derives the expectation and variance in the Neyman framework. Thus, it only requires the randomization-based finite sample inference and does not need Assumption 2.

LEMMA 1 (UNBIASEDNESS AND EXACT VARIANCE OF THE ESTIMATOR FOR THE SAPE) *Under Assumptions 1, 2, and 3, the expectation and variance of the estimator of the PAPE given in equation* (A4) *for estimating the SAPE defined in equation* (A3) *are given by,*

$$\begin{aligned}\mathbb{E}(\hat{\tau}_f^s \mid \mathcal{O}_n) &= \tau_f^s \\ \mathbb{V}(\hat{\tau}_f^s \mid \mathcal{O}_n) &= \frac{1}{n}\left(\frac{n_0}{n_1}\widetilde{S}_{f1}^2+\frac{n_1}{n_0}\widetilde{S}_{f0}^2+2\widetilde{S}_{f01}\right)\end{aligned}$$

*where* $\mathcal{O}_n=\{Y_i(1),Y_i(0),\mathbf{X}_i\}_{i=1}^n$ *and*

$$\widetilde{S}_{f01} = \frac{1}{n-1}\sum_{i=1}^n(\widetilde{Y}_{fi}(0)-\overline{\widetilde{Y}_{fi}(0)})(\widetilde{Y}_{fi}(1)-\overline{\widetilde{Y}_{fi}(1)}).$$

**Proof** We begin by computing the expectation with respect to the experimental treatment assignment, i.e., $T_i$,

$$\begin{aligned}
\mathbb{E}(\hat{\tau}_f^s \mid \mathcal{O}_n) &= \mathbb{E}\left\{\frac{1}{n_1}\sum_{i=1}^n f(\mathbf{X}_i)T_iY_i(1) + \frac{1}{n_0}\sum_{i=1}^n (1-f(\mathbf{X}_i))(1-T_i)Y_i(0)\right. \\
&\qquad \left. -\frac{\hat{p}_f}{n_1}\sum_{i=1}^n T_iY_i(1) - \frac{1-\hat{p}_f}{n_0}\sum_{i=1}^n (1-T_i)Y_i(0) \;\middle|\; \mathcal{O}_n\right\} \\
&= \frac{1}{n}\sum_{i=1}^n Y_i(1)f(\mathbf{X}_i) + \frac{1}{n}\sum_{i=1}^n Y_i(0)(1-f(\mathbf{X}_i)) - \frac{\hat{p}_f}{n}\sum_{i=1}^n Y_i(1) - \frac{1-\hat{p}_f}{n}\sum_{i=1}^n Y_i(0) \\
&= \tau_f^s
\end{aligned}$$

To derive the variance, we first rewrite the proposed estimator as,

$$\hat{\tau}_f^s = \tau_f^s + \sum_{i=1}^n D_i(f(\mathbf{X}_i) - \hat{p}_f)\left(\frac{Y_i(1)}{n_1} + \frac{Y_i(0)}{n_0}\right)$$

where $D_i = T_i - n_1/n$. Thus, noting $\mathbb{E}(D_i) = 0$, $\mathbb{E}(D_i^2) = n_0 n_1/n^2$, and $\mathbb{E}(D_iD_j) = -n_0n_1/\{n^2(n-1)\}$ for $i \neq j$, after some algebra, we have,

$$\begin{aligned}
\mathbb{V}(\hat{\tau}_f^s \mid \mathcal{O}_n) &= \mathbb{V}(\hat{\tau}_f^s - \tau_f^s \mid \mathcal{O}_n) = \mathbb{E}\left[\left\{\sum_{i=1}^n D_i\left(\frac{\widetilde{Y}_{fi}(1)}{n_1} + \frac{\widetilde{Y}_{fi}(0)}{n_0}\right)\right\}^2 \;\middle|\; \mathcal{O}_n\right] \\
&= \frac{1}{n}\left(\frac{n_0}{n_1}\widetilde{S}_{f1}^2 + \frac{n_1}{n_0}\widetilde{S}_{f0}^2 + 2\widetilde{S}_{f01}\right)
\end{aligned}$$

□

Now, we prove Theorem A2. Using Lemma 1 and the law of iterated expectation, we have,

$$\mathbb{E}(\hat{\tau}_f^s) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \{Y_i(f(\mathbf{X}_i)) - \hat{p}_fY_i(1) - (1-\hat{p}_f)Y_i(0)\}\right].$$

We compute the following expectation for $t = 0, 1$,

$$\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^n \hat{p}_fY_i(t)\right) &= \mathbb{E}\left(\sum_{i=1}^n \frac{\sum_{j=1}^n f(\mathbf{X}_j)}{n}Y_i(t)\right) = \frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^n f(\mathbf{X}_i)Y_i(t) + \sum_{i=1}^n\sum_{j\neq i} f(\mathbf{X}_j)Y_i(t)\right\} \\
&= \mathbb{E}\{f(\mathbf{X}_i)Y_i(t)\} + (n-1)p_f\mathbb{E}(Y_i(t)).
\end{aligned}$$

Putting them together yields the following bias expression,

$$\begin{aligned}
\mathbb{E}(\hat{\tau}_f^s) &= \mathbb{E}\left[\{Y_i(f(\mathbf{X}_i)) - \frac{1}{n}f(\mathbf{X}_i)\tau_i - \frac{n-1}{n}p_f\tau_i - Y_i(0)]\right] \\
&= \tau_f - \frac{1}{n}\mathbb{E}\left[\{f(\mathbf{X}_i)Y_i(1) - (1-f(\mathbf{X}_i))Y_i(0)\} - \{p_fY_i(1) - (1-p_f)Y_i(0)\}\right] \\
&= \tau_f - \frac{1}{n}\text{Cov}(f(\mathbf{X}_i), \tau_i).
\end{aligned}$$

where $\tau_i = Y_i(1) - Y_i(0)$. We can further rewrite the bias as,

$$
\begin{aligned}
-\frac{1}{n}\mathrm{Cov}(f(\mathbf{X}_i), \tau_i) &= \frac{1}{n} p_f\{\mathbb{E}(\tau_i \mid f(\mathbf{X}_i) = 1) - \tau\} \\
&= \frac{1}{n} p_f(1-p_f)\{\mathbb{E}(\tau_i \mid f(\mathbf{X}_i) = 1) - \mathbb{E}(\tau_i \mid f(\mathbf{X}_i) = 0)\} \\
&= \frac{\tau_f}{n}. \qquad \text{(A5)}
\end{aligned}
$$

where $\tau = \mathbb{E}(Y_i(1) - Y_i(0))$. This implies the estimator for the PAPE is unbiased, i.e., $\mathbb{E}(\hat{\tau}_f) = \tau_f$.

To derive the variance, Lemma 1 implies,

$$
\mathbb{V}(\hat{\tau}_f) = \frac{n^2}{(n-1)^2}\left[\mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n}\{\widetilde{Y}_{fi}(1) - \widetilde{Y}_{fi}(0)\}\right) + \mathbb{E}\left\{\frac{1}{n}\left(\frac{n_0}{n_1}\widetilde{S}_{f1}^2 + \frac{n_1}{n_0}\widetilde{S}_{f0}^2 + 2\widetilde{S}_{f01}\right)\right\}\right]. \qquad \text{(A6)}
$$

Applying Lemma 1 of Nadeau and Bengio (2000) to the first term within the square brackets yields,

$$
\mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n}\{\widetilde{Y}_{fi}(1) - \widetilde{Y}_{fi}(0)\}\right) = \mathrm{Cov}(\widetilde{Y}_{fi}(1) - \widetilde{Y}_{fi}(0), \widetilde{Y}_{fi}(1) - \widetilde{Y}_{fi}(0)) + \frac{1}{n}\mathbb{E}(\widetilde{S}_{f1}^2 + \widetilde{S}_{f0}^2 - 2\widetilde{S}_{f01}), \qquad \text{(A7)}
$$

where $i \neq j$. Focusing on the covariance term, we have,

$$
\begin{aligned}
& \mathrm{Cov}(\widetilde{Y}_{fi}(1) - \widetilde{Y}_{fi}(0), \widetilde{Y}_{fi}(1) - \widetilde{Y}_{fi}(0)) \\
= \ & \mathrm{Cov}\left(\left\{f(\mathbf{X}_i) - \frac{1}{n}\sum_{i'=1}^{n} f(\mathbf{X}_{i'})\right\}\tau_i, \left\{f(\mathbf{X}_j) - \frac{1}{n}\sum_{j'=1}^{n} f(\mathbf{X}_{j'})\right\}\tau_j\right) \\
= \ & -2\mathrm{Cov}\left(\frac{n-1}{n} f(\mathbf{X}_i)\tau_i, \frac{1}{n} f(\mathbf{X}_i)\tau_j\right) + \sum_{i' \neq i,j} \mathrm{Cov}\left(\frac{1}{n} f(\mathbf{X}_{i'})\tau_i, \frac{1}{n} f(\mathbf{X}_{i'})\tau_j\right) \\
& + 2\sum_{i' \neq i,j} \mathrm{Cov}\left(\frac{1}{n} f(\mathbf{X}_j)\tau_i, \frac{1}{n} f(\mathbf{X}_{i'})\tau_j\right) + \mathrm{Cov}\left(\frac{1}{n} f(\mathbf{X}_j)\tau_i, \frac{1}{n} f(\mathbf{X}_i)\tau_j\right) \\
= \ & -\frac{2(n-1)\tau}{n^2}\mathrm{Cov}\left(f(\mathbf{X}_i), f(\mathbf{X}_i)\tau_i\right) + \frac{(n-2)\tau^2}{n^2}\mathbb{V}(f(\mathbf{X}_i)) \\
& + \frac{2(n-2)\tau}{n^2} p_f \mathrm{Cov}\left(f(\mathbf{X}_i), \tau_i\right) + \frac{1}{n^2}\left\{\mathrm{Cov}^2\left(f(\mathbf{X}_i), \tau_i\right) + 2p_f\tau\mathrm{Cov}\left(f(\mathbf{X}_i), \tau_i\right)\right\} \\
= \ & \frac{1}{n^2}\mathrm{Cov}^2\left(f(\mathbf{X}_i), \tau_i\right) + \frac{(n-2)\tau^2}{n^2} p_f(1-p_f) + \frac{2(n-1)\tau}{n^2}\mathrm{Cov}\left(f(\mathbf{X}_i), (p_f - f(\mathbf{X}_i))\tau_i\right) \\
= \ & \frac{1}{n^2}\left[\tau_f^2 + (n-2)p_f(1-p_f)\tau^2 + 2(n-1)\tau\left\{p_f\tau_f - (1-p_f)\mathbb{E}(f(\mathbf{X}_i)\tau_i)\right\}\right] \\
= \ & \frac{1}{n^2}\left[\tau_f^2 + (n-2)p_f(1-p_f)\tau^2 + 2(n-1)\tau\left\{\tau_f(2p_f - 1) - (1-p_f)p_f\tau\right\}\right] \\
= \ & \frac{1}{n^2}\left\{\tau_f^2 - np_f(1-p_f)\tau^2 + 2(n-1)(2p_f-1)\tau_f\tau\right\},
\end{aligned}
$$

where the third equality follows from the formula for the covariance of products of two random variables (Bohrnstedt and Goldberger, 1969). Finally, combining this result with

| Individual | $T_i$ | $f(\mathbf{X}_i)$ | $Y_i$ | $Y_i(0)$ | $Y_i(1)$ |
|---|---|---|---|---|---|
| A | 1 | 1 | 2 | 0 | 2 |
| B | 1 | 0 | 3 | 1 | 3 |
| C | 0 | 0 | $-1$ | $-1$ | $-1$ |
| D | 0 | 1 | 1 | 1 | 0 |
| E | 1 | 0 | 3 | 0 | 3 |

Table A1: A Numerical Example for Binary Treatment Assignment and Outcomes

equations (A6) and (A7) yields,

$$\mathbb{V}(\hat{\tau}_f) = \frac{n^2}{(n-1)^2}\left[\frac{\mathbb{E}(\widetilde{S}_{f1}^2)}{n_1} + \frac{E(\widetilde{S}_{f0}^2)}{n_0} + \frac{1}{n^2}\left\{\tau_f^2 - np_f(1-p_f)\tau^2 + 2(n-1)(2p_f-1)\tau_f\tau\right\}\right].$$

□

A potential complication with this estimator in practice is that its estimate (along with the variance) would change under an additive transformation, i.e., $Y_i(t) \to Y_i(t) + \delta$ for $t = 0, 1$ and a given constant $\delta$. This issue is not due to the specific construction of the proposed estimator. It instead reflects the fundamental issue of many prescription effects including the population average value and PAPE that they cannot be defined solely in terms of multiples of $Y_i(1) - Y_i(0)$ (see Appendix A.1.3 for a numerical example). One solution is to center the outcome variable such that $\sum_{i=1}^n Y_i T_i / n_1 + \sum_{i=1}^n Y_i(1-T_i)/n_0 = 0$ holds. This solution is motivated by the fact that when the condition holds in the population (i.e. $\mathbb{E}(Y_i(1) + Y_i(0)) = 0$), the variance of the PAPE estimator is minimized.

### A.1.3 A Numerical Example Showing the Lack of Additive Invariance for the Population Average Value

Consider an ITR $f : \mathcal{X} \to \{0, 1\}$, and we would like to know its population average value. Table A1 shows an numerical example with the observed outcome $Y_i$, the ITR $f(X_i)$, the actual assignment $T_i$, and the potential outcomes $Y_i(0), Y_i(1)$. Then, in this example, we have $n_1 = 3$ (A,B,E), $n_0 = 2$ (C,D), and the population average value estimator would be:

$$\begin{aligned}
\hat{\lambda}_f(\mathcal{Z}) &= \frac{1}{n_1}\sum_{i=1}^n Y_i T_i f(\mathbf{X}_i) + \frac{1}{n_0}\sum_{i=1}^n Y_i(1-T_i)(1-f(\mathbf{X}_i)) \\
&= \frac{1}{3}(1\cdot 2 + 0\cdot 3 + 0 \cdot 3) + \frac{1}{2}(1\cdot -1 + 0\cdot 1) \\
&= \frac{1}{6}
\end{aligned}$$

Now let us consider an additive transformation $Y_i(t) \to Y_i(t) + 1 := Y_i'(t)$ for $t = 0, 1$, where every outcome value is raised by 1. Then, its population average value estimator is now:

$$\begin{aligned}
\hat{\lambda}_f'(\mathcal{Z}) &= \frac{1}{n_1}\sum_{i=1}^n Y_i' T_i f(\mathbf{X}_i) + \frac{1}{n_0}\sum_{i=1}^n Y_i'(1-T_i)(1-f(\mathbf{X}_i)) \\
&= \frac{1}{3}(1\cdot 3 + 0\cdot 4 + 0\cdot 4) + \frac{1}{2}(1\cdot 0 + 0\cdot 2)
\end{aligned}$$

$$= 1$$

Note that the difference $\hat{\lambda}'_f(\mathcal{Z}) - \hat{\lambda}_f(\mathcal{Z}) = \frac{5}{6} \neq 1$ does not equal to the amount of additive transformation. The problem arises because they are not multiples of $Y_i(1) - Y_i(0)$ but rather they depend on what the actual assignments of the ITR.

## A.2 Proof of Theorem 1

We begin by deriving the variance. The derivation proceeds in the same fashion as the one for Theorem A2. The main difference lies in the derivation of the covariance term, which we detail below. First, we note that,

$$\begin{aligned}
\Pr(f(\mathbf{X}_i, \hat{c}_p(f)) = 1) &= \int_{-\infty}^{\infty} \Pr(f(\mathbf{X}_i, c) = 1 \mid \hat{c}_p(f) = c) P(\hat{c}_p(f) = c) \mathrm{d}c \\
&= \int_{-\infty}^{\infty} \frac{\lfloor np \rfloor}{n} P(\hat{c}_p(f) = c) \mathrm{d}c \\
&= \frac{\lfloor np \rfloor}{n},
\end{aligned}$$

where the second equality follows from the fact that once conditioned on $\hat{c}_p(f) = c$, exactly $\lfloor np \rfloor$ out of $n$ units will be assigned to the treatment condition. Given this result, we can compute the covariance as follows,

$$\begin{aligned}
& \mathrm{Cov}(\widetilde{Y}_i(1) - \widetilde{Y}_i(0), \widetilde{Y}_j(1) - \widetilde{Y}_j(0)) \\
= & \mathrm{Cov}\left\{ (f(\mathbf{X}_i, \hat{c}_p(f)) - p)\, \tau_i, (f(\mathbf{X}_j, \hat{c}_p(f)) - p)\, \tau_j \right\} \\
= & \mathrm{Cov}\left\{ f(\mathbf{X}_i, \hat{c}_p(f))\tau_i, f(\mathbf{X}_j, \hat{c}_p(f))\tau_j \right\} - 2p\mathrm{Cov}\left(\tau_i, f(\mathbf{X}_j, \hat{c}_p(f))\tau_j\right) \\
= & \frac{n\lfloor np \rfloor(\lfloor np \rfloor - 1) - \lfloor np \rfloor^2 (n-1)}{n^2(n-1)} \mathbb{E}(\tau_i \mid f(\mathbf{X}_i, \hat{c}_p(f)) = 1)^2 - 2p\mathrm{Cov}\left(\tau_i, f(\mathbf{X}_j, \hat{c}_p(f))\tau_j\right) \\
= & \frac{\lfloor np \rfloor(\lfloor np \rfloor - n)}{n^2(n-1)} \kappa_1(\hat{c}_p(f))^2 + \frac{2p\lfloor np \rfloor(n - \lfloor np \rfloor)}{n^2(n-1)} \left(\kappa_1(\hat{c}_p(f))^2 - \kappa_1(\hat{c}_p(f))\kappa_0(\hat{c}_p(f))\right) \\
= & (2p-1)\frac{\lfloor np \rfloor(n - \lfloor np \rfloor)}{n^2(n-1)} \kappa_1(\hat{c}_p(f))^2 - \frac{2p\lfloor np \rfloor(n - \lfloor np \rfloor)}{n^2(n-1)} \kappa_1(\hat{c}_p(f))\kappa_0(\hat{c}_p(f)) \\
= & \frac{\lfloor np \rfloor(n - \lfloor np \rfloor)}{n^2(n-1)} \left\{ (2p-1)\kappa_1(\hat{c}_p(f))^2 - 2p\kappa_1(\hat{c}_p(f))\kappa_0(\hat{c}_p(f)) \right\}.
\end{aligned}$$

Combining this covariance result with the expression for the marginal variances yields the desired variance expression for $\hat{\tau}_f(\hat{c}_p(f))$.

Next, we derive the upper bound of bias. Using the same technique as the proof of Theorem A2, we can rewrite the expectation of the proposed estimator as,

$$\mathbb{E}(\hat{\tau}_f(c_p)) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \left\{ Y_i\left(f(\mathbf{X}_i, \hat{c}_p(f))\right) - pY_i(1) - (1-p)Y_i(0) \right\}\right].$$

Now, define $F(c) = \mathbb{P}(s(\mathbf{X}_i) \leq c)$. Without loss of generality, assume $\hat{c}_p(f) > c_p$ (If this is not the case, we simply switch the upper and lower limits of the integrals below). Then, the bias of the estimator is given by,

$$|\mathbb{E}(\hat{\tau}_f(\hat{c}_p(f))) - \tau_f(c_p)| = \left| \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n \left\{ Y_i\left(f(\mathbf{X}_i, \hat{c}_p(f))\right) - Y_i\left(f(\mathbf{X}_i, c_p)\right) \right\}\right] \right|$$

$$
\begin{aligned}
&= \left| \mathbb{E}_{\hat{c}_p(f)} \left[ \int_{c_p}^{\hat{c}_p(f)} \mathbb{E}(\tau_i \mid s(\mathbf{X}_i) = c) \mathrm{d}F(c) \right] \right| \\
&= \left| \mathbb{E}_{F(\hat{c}_p(f))} \left[ \int_{F(c_p)}^{F(\hat{c}_p(f))} \mathbb{E}(\tau_i \mid s(\mathbf{X}_i) = F^{-1}(x)) \mathrm{d}x \right] \right| \\
&\leq \mathbb{E}_{F(\hat{c}_p(f))} \left[ |F(\hat{c}_p(f)) - (1-p)| \times \max_{c \in [c_p, \hat{c}_p(f)]} |\mathbb{E}(\tau_i \mid s(\mathbf{X}_i) = c)| \right].
\end{aligned}
$$

By the definition of $\hat{c}_p(f)$, $F(\hat{c}_p(f))$ is the $(n - \lfloor np \rfloor)$th order statistic of $n$ independent uniform random variables, and thus follows the Beta distribution with the shape and scale parameters equal to $n - \lfloor np \rfloor$ and $\lfloor np \rfloor + 1$, respectively. For the special case where $p = 1$, we define the 0th order statistic of $n$ uniform random variables to be 0, and by extension also define the "beta distribution" with shape parameter $\leq 0$ to be $H(x)$ where $H(x)$ is the Heaviside step function. Therefore, we have,

$$
\mathbb{P}(|F(\hat{c}_p(f)) - p| > \epsilon) = 1 - B(1-p+\epsilon, n-\lfloor np \rfloor, \lfloor np \rfloor + 1) + B(1-p-\epsilon, n-\lfloor np \rfloor, \lfloor np \rfloor + 1), \tag{A8}
$$

where $B(\epsilon, \alpha, \beta)$ is the incomplete beta function, i.e.,

$$
B(\epsilon, \alpha, \beta) = \int_0^{\epsilon} t^{\alpha-1}(1-t)^{\beta-1} \mathrm{d}t.
$$

Combining with the result above, the desired result follows. □

## A.3 Estimation and Inference of the Population Average Prescriptive Difference of Fixed ITRs

THEOREM A3 (BIAS AND VARIANCE OF THE PAPD ESTIMATOR WITH A BUDGET CONSTRAINT) *Under Assumptions 1, 2, and 3, the bias of the proposed estimator of the PAPD with a budget constraint* p *defined in equation* (8) *can be bounded as follows,*

$$
\begin{aligned}
\mathbb{P}_{\hat{c}_p(f), \hat{c}_p(g)}(|\mathbb{E}\{\widehat{\Delta}_p(f, g, \mathcal{Z}) - \Delta_p(f, g) \mid \hat{c}_p(f), \hat{c}_p(g)\}| \geq \epsilon) &\leq 1 - 2B(1 - p + \gamma_p(\epsilon), n - \lfloor np \rfloor, \lfloor np \rfloor + 1) \\
&\quad + 2B(1 - p - \gamma_p(\epsilon), n - \lfloor np \rfloor, \lfloor np \rfloor + 1),
\end{aligned}
$$

*where any given constant $\epsilon > 0$, $B(\epsilon, \alpha, \beta)$ is the incomplete beta function (if $\alpha = 0$ and $\beta > 0$, we set $B(\epsilon, \alpha, \beta) := H(\epsilon)$ for all $\epsilon$ where $H(\epsilon)$ is the Heaviside step function), and*

$$
\gamma_p(\epsilon) = \frac{\epsilon}{\max_{c \in [c_p(f)-\epsilon,\ c_p(f)+\epsilon],\ d \in [c_p(g)-\epsilon,\ c_p(g)+\epsilon]} \{\mathbb{E}(\tau_i \mid s_f(\mathbf{X}_i) = c), \mathbb{E}(\tau_i \mid s_g(\mathbf{X}_i) = d)\}}.
$$

*The variance of the estimator is,*

$$
\begin{aligned}
\mathbb{V}(\widehat{\Delta}_p(f, g, \mathcal{Z})) &= \frac{\mathbb{E}(S_{fgp1}^2)}{n_1} + \frac{\mathbb{E}(S_{fgp0}^2)}{n_0} + \frac{\lfloor np \rfloor(\lfloor np \rfloor - n)}{n^2(n-1)}(\kappa_{f1}(p)^2 + \kappa_{g1}(p)^2) \\
&\quad -2\left(\Pr(f(\mathbf{X}_i, \hat{c}_p(f)) = g(\mathbf{X}_i, \hat{c}_p(g)) = 1) - \frac{\lfloor np \rfloor^2}{n^2}\right)\kappa_{f1}(p)\kappa_{g1}(p),
\end{aligned}
$$

*where $S_{fgpt}^2 = \sum_{i=1}^n (Y_{fgpi}(t) - \overline{Y_{fgp}(t)})^2/(n-1)$ with $Y_{fgpi}(t) = \{f(\mathbf{X}_i, \hat{c}_p(f)) - g(\mathbf{X}_i, \hat{c}_p(g))\} Y_i(t)$ and $\overline{Y_{fgp}(t)} = \sum_{i=1}^n Y_{fgpi}(t)/n$ for $t = 0, 1$.*

To estimate the variance, it is tempting to replace all the unknown parameters with their sample analogues. However, unlike the case of the variance of the PAPE estimator under a budget constraint (see Theorem 1), there is no useful identity for the joint probability $\Pr(f(\mathbf{X}_i, \hat{c}_p(f)) = g(\mathbf{X}_i, \hat{c}_p(g)) = 1)$ under general $g$. Thus, an empirical analogue of $\hat{c}_p(f)$ and $\hat{c}_p(g)$ is not a good estimate because it is solely based on one realization. Thus, we use the following conservative bound,

$$\begin{aligned} & -\left(\Pr(f(\mathbf{X}_i, \hat{c}_p(f)) = g(\mathbf{X}_i, \hat{c}_p(g)) = 1) - \frac{\lfloor np \rfloor^2}{n^2}\right) \kappa_{f1}(p)\kappa_{g1}(p) \\ \leq \quad & \frac{\lfloor np \rfloor \max\{\lfloor np \rfloor, n - \lfloor np \rfloor\}}{n^2(n-1)} |\kappa_{f1}(p)\kappa_{g1}(p)|, \end{aligned}$$

where the inequality follows because the maximum is achieved when the scoring rules of $f$ and $g$, i.e., $s_f(\mathbf{X}_i)$ and $s_g(\mathbf{X}_i)$, are perfectly correlated. We use this upper bound in our simulation and empirical studies. In Section 5, we find that this upper bound estimate of the variance produces only a small conservative bias.

**Proof** The proof of the bounds for the expectation and variance of the proposed estimator largely follows the proof given in Appendix A.2. The only significant difference is the calculation of the covariance term, which is given below.

$$\begin{aligned} & \operatorname{Cov}(Y_i^*(1) - Y_i^*(0), Y_j^*(1) - Y_j^*(0)) \\ = \quad & \operatorname{Cov}\left(\{f(\mathbf{X}_i, \hat{c}_p(f)) - g(\mathbf{X}_i, \hat{c}_p(g))\}\tau_i, \{f(\mathbf{X}_j, \hat{c}_p(f)) - g(\mathbf{X}_j, \hat{c}_p(g))\}\tau_j\right) \\ = \quad & \operatorname{Cov}(f(\mathbf{X}_i, \hat{c}_p(f))\tau_i, f(\mathbf{X}_j, \hat{c}_p(f))\tau_j) + \operatorname{Cov}(g(\mathbf{X}_i, \hat{c}_p(g))\tau_i, g(\mathbf{X}_j, \hat{c}_p(g))\tau_j) \\ & -2\operatorname{Cov}(f(\mathbf{X}_i, \hat{c}_p(f))\tau_i, g(\mathbf{X}_j, \hat{c}_p(g))\tau_j) \\ = \quad & \frac{\lfloor np \rfloor(\lfloor np \rfloor - n)}{n^2(n-1)}\left(\kappa_{f1}(p)^2 + \kappa_{g1}(p)^2\right) - 2\operatorname{Cov}(f(\mathbf{X}_i, \hat{c}_p(f))\tau_i, g(\mathbf{X}_j, \hat{c}_p(g))\tau_j) \\ = \quad & \frac{\lfloor np \rfloor(\lfloor np \rfloor - n)}{n^2(n-1)}\left(\kappa_{f1}(p)^2 + \kappa_{g1}(p)^2\right) \\ & -2\left(\Pr(f(\mathbf{X}_i, \hat{c}_p(f)) = g(\mathbf{X}_i, \hat{c}_p(g)) = 1) - \frac{\lfloor np \rfloor^2}{n^2}\right)\kappa_{f1}(p)\kappa_{g1}(p) \end{aligned}$$

□

## A.4 Proof of Theorem 2

The derivation of the variance expression in Theorem 2 proceeds in the same fashion as Theorem A2 (see Appendix A.2) with the only non-trivial change being the calculation of the covariance term. Note $\Pr(f(\mathbf{X}_i, \hat{c}_{\frac{k}{n}}(f)) = 1) = k/n$ for $t = 0, 1$ and $n_f = Z \sim \text{Binom}(n, p_f)$. Then, we have:

$$\begin{aligned} & \operatorname{Cov}(Y_i^*(1) - Y_i^*(0), Y_j^*(1) - Y_j^*(0)) \\ = \quad & \operatorname{Cov}\left[\left\{\frac{1}{n}\left(\sum_{k=1}^{n_f} f(\mathbf{X}_i, \hat{c}_{k/n}(f)) + \sum_{k=n_f+1}^{n} f(\mathbf{X}_i, \hat{c}_{n_f/n}(f))\right) - \frac{1}{2}\right\}\tau_i,\right. \end{aligned}$$

$$
\left\{\frac{1}{n}\left(\sum_{k=1}^{n_f} f(\mathbf{X}_j, \hat{c}_{k/n}(f)) + \sum_{k=n_f+1}^{n} f(\mathbf{X}_j, \hat{c}_{n_f/n}(f))\right) - \frac{1}{2}\right\}\tau_j\Bigg]
$$

$$
\begin{aligned}
= \;& \mathbb{E}\Bigg\{\mathrm{Cov}\Bigg[\left\{\frac{1}{n}\left(\sum_{k=1}^{Z} f(\mathbf{X}_i, \hat{c}_{k/n}(f)) + \sum_{k=Z+1}^{n} f(\mathbf{X}_i, \hat{c}_{Z/n}(f))\right) - \frac{1}{2}\right\}\tau_i, \\
& \left\{\frac{1}{n}\left(\sum_{k=1}^{Z} f(\mathbf{X}_j, \hat{c}_{k/n}(f)) + \sum_{k=Z+1}^{n} f(\mathbf{X}_j, \hat{c}_{Z/n}(f))\right) - \frac{1}{2}\right\}\tau_j \;\Big|\; Z\Bigg]\Bigg\} \\
& + \mathrm{Cov}\Bigg\{\mathbb{E}\Bigg[\left\{\frac{1}{n}\left(\sum_{k=1}^{Z} f(\mathbf{X}_i, \hat{c}_{k/n}(f)) + \sum_{k=Z+1}^{n} f(\mathbf{X}_i, \hat{c}_{Z/n}(f))\right) - \frac{1}{2}\right\}\tau_i \;\Big|\; Z\Bigg], \\
& \mathbb{E}\Bigg[\left\{\frac{1}{n}\left(\sum_{k=1}^{Z} f(\mathbf{X}_j, \hat{c}_{k/n}(f)) + \sum_{k=Z+1}^{n} f(\mathbf{X}_j, \hat{c}_{Z/n}(f))\right) - \frac{1}{2}\right\}\tau_j \;\Big|\; Z\Bigg]\Bigg\} \\
= \;& \mathbb{E}\Bigg[-\frac{1}{n}\left\{\sum_{k=1}^{Z}\frac{k(n-k)}{n^2(n-1)}\kappa_{f1}(k/n)\kappa_{f0}(k/n) + \frac{Z(n-Z)^2}{n^2(n-1)}\kappa_{f1}(Z/n)\kappa_{f0}(Z/n)\right\} \\
& -\frac{2}{n^4(n-1)}\sum_{k=1}^{Z-1}\sum_{k'=k+1}^{Z} k(n-k')\kappa_{f1}(k/n)\kappa_{f1}(k'/n) \\
& -\frac{Z^2(n-Z)^2}{n^4(n-1)}\kappa_{f1}(Z/n)^2 - \frac{2(n-Z)^2}{n^4(n-1)}\sum_{k=1}^{Z} k\kappa_{f1}(Z/n)\kappa_{f1}(k/n) \\
& +\frac{1}{n^4}\sum_{k=1}^{Z} k(n-k)\kappa_{f1}(k/n)^2\Bigg] + \mathbb{V}\left(\sum_{i=1}^{Z}\frac{i}{n}\kappa_{f1}(i/n) + \frac{(n-Z)Z}{n}\kappa_{f1}(Z/n)\right),
\end{aligned}
$$

where the last equality is based on the results from Appendix A.2.

For the bias, we can rewrite $\Gamma_f$ as,

$$
\Gamma_f \;=\; \int_0^{p_f}\mathbb{E}\{Y_i(f(\mathbf{X}_i, c_p))\}\mathrm{d}p + (1-p_f)\mathbb{E}\{Y_i(f(\mathbf{X}_i, c^*))\} - \frac{1}{2}\mathbb{E}\{Y_i(1)+Y_i(0)\},
$$

and similarly its estimator $\widehat{\Gamma}_f$ as,

$$
\mathbb{E}(\widehat{\Gamma}_f) \;=\; \mathbb{E}\left\{\int_0^{\hat{p}_f} Y_i(f(\mathbf{X}_i, c_p))\mathrm{d}p\right\} + \mathbb{E}\{(1-\hat{p}_f)Y_i(f(\mathbf{X}_i, c^*))\} - \frac{1}{2}\mathbb{E}\{Y_i(1)+Y_i(0)\}.
$$

Therefore, the bias of the estimator is, using a derivation similar to Appendix A.2:

$$
\begin{aligned}
\left|\mathbb{E}(\widehat{\Gamma}_f) - \Gamma_f\right| \;\leq\;& \mathbb{E}\Bigg[|p_f - \hat{p}_f| \max_{c\in\{\min\{\hat{p}_f, p_f\}, \max\{\hat{p}_f, p_f\}\}} \mathbb{E}\{Y_i(f(\mathbf{X}_i, c)) - Y_i(f(\mathbf{X}_i, c^*))\}| \\
& + |\mathbb{E}\{Y_i(f(\mathbf{X}_i, c^*))\} - \mathbb{E}\{Y_i(f(\mathbf{X}_i, \hat{c}_{p_f}))\}|\Bigg] \\
\leq\;& (\epsilon+1)\max_{c\in[c^*-\epsilon, c^*+\epsilon]} |\mathbb{E}[Y_i(f(\mathbf{X}_i, c)) - Y_i(f(\mathbf{X}_i, c^*))]| \\
\leq\;& (\epsilon+1)\epsilon \max_{c\in[c^*-\epsilon, c^*+\epsilon]} |\mathbb{E}(\tau_i \mid s(\mathbf{X}_i) = c)|\,.
\end{aligned}
$$

Now, taking the bound $\epsilon(1+\epsilon) \leq 2\epsilon$ for $0 \leq \epsilon \leq 1$ in equation (A8) of Appendix A.2, we have the desired result. □

## A.5 Evaluation of an Estimated ITR with No Budget Constraint

Formally, we define an machine learning algorithm $F$ to be a deterministic map from the space of observable data $\mathcal{Z} = \{\mathcal{X}, \mathcal{T}, \mathcal{Y}\}$ to the space of ITRs $\mathcal{F}$,

$$F : \mathcal{Z} \longrightarrow \mathcal{F}.$$

We emphasize that no restriction is placed on the machine learning algorithm $F$ or ITR $f$.

To extend the population average value (equation (1)), we consider the average performance of an estimated ITR across training data sets of fixed size. First, for any given values of pre-treatment variables $\mathbf{X}_i = \mathbf{x}$, we define the average treatment proportion under the estimated ITR obtained by applying the machine learning algorithm $F$ to training data $\mathcal{Z}^{tr}$ of size $n - m$,

$$\bar{f}_F(\mathbf{x}) \ = \ \mathbb{E}\{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{x}) \mid \mathbf{X}_i = \mathbf{x}\} \ = \ \Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{x}) = 1 \mid \mathbf{X}_i = \mathbf{x})$$

where the expectation is taken over the random sampling of training data $\mathcal{Z}^{tr}$. Although $\bar{f}_F$ depends on the training data size, we suppress it to ease notational burden.

Then, the population average value of an estimated ITR can be defined as,

$$\lambda_F \ = \ \mathbb{E}\left\{\bar{f}_F(\mathbf{X}_i)Y_i(1) + (1 - \bar{f}_F(\mathbf{X}_i))Y_i(0)\right\} \tag{A9}$$

where the expectation is taken over the population distribution of $\{\mathbf{X}_i, Y_i(1), Y_i(0)\}$. In contrast to the population average value of a fixed ITR, this estimand accounts for the estimation uncertainty of the ITR by averaging over the random sampling of training sets.

To generalize the PAPE (equation (2)), we first define the population proportion of units assigned to the treatment condition under the estimated ITR as follows,

$$p_F \ = \ \mathbb{E}\{\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = 1 \mid \mathbf{X}_i)\} \ = \ \mathbb{E}\{\bar{f}_F(\mathbf{X}_i)\}$$

where the expectation is taken with respect to the sampling of training data of size $n - m$ and the population distribution of $\mathbf{X}_i$. Then, the PAPE of an estimated ITR is given by,

$$\tau_F \ = \ \mathbb{E}\{\lambda_F - p_F Y_i(1) - (1 - p_F)Y_i(0)\}, \tag{A10}$$

where $\lambda_F$ is the population average value of the estimated ITR defined in equation (A9).

### A.5.1 The Population Average Value

We begin by considering the following cross-validation estimator of the population average value for an estimated ITR (equation (A9)),

$$\hat{\lambda}_F \ = \ \frac{1}{K}\sum_{k=1}^{K} \hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k). \tag{A11}$$

The following theorem proves the unbiasedness of this estimator and derives its exact variance expression under the Neyman's repeated sampling framework.

THEOREM A4 (UNBIASEDNESS AND EXACT VARIANCE OF THE CROSS-VALIDATION POPULATION AVERAGE VALUE ESTIMATOR) *Under Assumptions 1, 2, and 3, the expectation and variance of the cross-validation Population Average Value estimator defined in equation* (A11) *are given by,*

$$\begin{aligned}
\mathbb{E}(\hat{\lambda}_F) &= \lambda_F \\
\mathbb{V}(\hat{\lambda}_F) &= \frac{\mathbb{E}(S^2_{\hat{f}1})}{m_1} + \frac{\mathbb{E}(S^2_{\hat{f}0})}{m_0} + \mathbb{E}\left\{\text{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right\} - \frac{K-1}{K}\mathbb{E}(S^2_F)
\end{aligned}$$

*for* $i \neq j$ *where* $S^2_{\hat{f}t} = \sum_{i=1}^m \left(Y_{\hat{f}i}(t) - \overline{Y_{\hat{f}}(t)}\right)^2/(m-1)$, $S^2_F = \sum_{k=1}^K \left\{\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k) - \overline{\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k)}\right\}^2/(K-1)$, *and* $\tau_i = Y_i(1) - Y_i(0)$ *with* $Y_{\hat{f}i}(t) = \mathbf{1}\{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = t\}Y_i(t)$, $\overline{Y_{\hat{f}}(t)} = \sum_{i=1}^m Y_{\hat{f}i}(t)/m$, *and* $\overline{\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k)} = \sum_{k=1}^K \hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k)/K$ *for* $t = \{0, 1\}$.

The proof of unbiasedness is similar to that of Appendix A.1.2 and thus is omitted. To derive the variance, we first introduce the following useful lemma, adapted from Nadeau and Bengio (2000).

LEMMA 2

$$\begin{aligned}
\mathbb{E}(S^2_F) &= \mathbb{V}(\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k)) - \text{Cov}(\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k), \hat{\lambda}_{\hat{f}_{-\ell}}(\mathcal{Z}_\ell)), \\
\mathbb{V}(\hat{\lambda}_F) &= \frac{\mathbb{V}(\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k))}{K} + \frac{K-1}{K}\text{Cov}(\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k), \hat{\lambda}_{\hat{f}_{-\ell}}(\mathcal{Z}_\ell)).
\end{aligned}$$

*where* $k \neq \ell$.

The lemma implies,

$$\mathbb{V}(\hat{\lambda}_F) = \mathbb{V}(\hat{\lambda}_{f_{-k}}(\mathcal{Z}_k)) - \frac{K-1}{K}\mathbb{E}(S^2_F). \tag{A12}$$

We then follow the same process of derivation as in Appendix A.1.2 while replacing $Y_i^*(t)$ with $Y_{\hat{f}i}(t)$ for $t \in \{0, 1\}$. The only difference lies in the covariance term, which can be expanded as follows,

$$\begin{aligned}
\text{Cov}(Y_{\hat{f}i}(1) - Y_{\hat{f}i}(0), Y_{\hat{f}j}(1) - Y_{\hat{f}j}(0)) &= \text{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i + Y_i(0), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\tau_j + Y_j(0)) \\
&= \text{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\tau_j) \\
&= \mathbb{E}\left[\text{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\tau_j \mid \mathbf{X}_i, \mathbf{X}_j, \tau_i, \tau_j)\right] \\
&= \mathbb{E}\left[\text{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right].
\end{aligned}$$

So the full variance expression is:

$$\begin{aligned}
\mathbb{V}(\hat{\lambda}_F) &= \mathbb{V}(\hat{\lambda}_{f_{-k}}(\mathcal{Z}_k)) - \frac{K-1}{K}\mathbb{E}(S^2_F) \\
&= \frac{\mathbb{E}(S^2_{\hat{f}1})}{m_1} + \frac{\mathbb{E}(S^2_{\hat{f}0})}{m_0} + \mathbb{E}\left[\text{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right] - \frac{K-1}{K}\mathbb{E}(S^2_F)
\end{aligned}$$

□

When compared to the case of a fixed ITR with the sample size of $m$ (see Theorem A1 in Appendix A.1.1), the variance has two additional terms. The covariance term accounts for the estimation uncertainty about the ITR, and is typically positive because two units, for which an estimated ITR makes the same treatment assignment recommendation, are likely to have causal effects with the same sign. The second term is due to the efficiency gain resulting from the $K$-fold cross-validation rather than evaluating an estimated ITR once.

The cross-validation estimate of $\mathbb{E}(S^2_{\hat{f}t})$ is straightforward and is given by,

$$\widehat{\mathbb{E}(S^2_{\hat{f}t})} \;=\; \frac{1}{K(m_t-1)}\sum_{k=1}^{K}\sum_{i=1}^{m}\mathbf{1}\{T_i^{(k)}=t\}\left\{Y_{\hat{f}i}^{(k)}(t)-\overline{Y_{\hat{f}t}^{(k)}}\right\}^2,$$

where $Y_{\hat{f}i}^{(k)}(t) = \mathbf{1}\{\hat{f}_{-k}(\mathbf{X}_i^{(k)}) = t\}Y_i^{(k)}(t)$ and $\overline{Y_{\hat{f}t}^{(k)}} = \sum_{i=1}^{m}\mathbf{1}\{T_i^{(k)} = t\}Y_{\hat{f}i}^{(k)}(t)/m_t$. In contrast, the estimation of this cross-validation variance requires care. In particular, although it is tempting to estimate $\mathbb{E}(S_F^2)$ using the realization of $S_F^2$, this estimate is highly variable especially when $K$ is small. As a result, it often yields a negative overall variance estimate. We address this problem by first noting that Lemma 2 implies,

$$\mathbb{V}(\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k)) \;=\; \frac{\mathbb{E}(S^2_{\hat{f}1})}{m_1}+\frac{\mathbb{E}(S^2_{\hat{f}0})}{m_0}+\mathbb{E}\left\{\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i),\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\mid\mathbf{X}_i,\mathbf{X}_j)\tau_i\tau_j\right\} \;\geq\; \mathbb{E}(S_F^2).$$

Then, this inequality suggests the following consistent estimator of $\mathbb{E}(S_F^2)$,

$$\widehat{\mathbb{E}(S_F^2)} \;=\; \min\left(S_F^2, \mathbb{V}(\widehat{\hat{\lambda}_{\hat{f}_{-k}}(\mathcal{Z}_k)})\right).$$

Although this yields a conservative estimate of $\mathbb{V}(\hat{\lambda}_F)$ in finite samples, the bias appears to be small in practice (see Section 5.2).

Finally, for the estimation of the covariance term, since $\hat{f}_{\mathcal{Z}^{tr}}$ is binary, we have,

$$\begin{aligned}&\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i),\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\mid\mathbf{X}_i,\mathbf{X}_j)\\ =\;&\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)=\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)=1\mid\mathbf{X}_i,\mathbf{X}_j)-\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)=1\mid\mathbf{X}_i)\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)=1\mid\mathbf{X}_j),\end{aligned}$$

for $i \neq j$. An unbiased cross-validation estimator of this covariance (given $\mathbf{X}_i$ and $\mathbf{X}_j$) is,

$$\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i),\widehat{\hat{f}_{\mathcal{Z}^{tr}}}(\mathbf{X}_j)\mid\mathbf{X}_i,\mathbf{X}_j) \;=\; \frac{1}{K}\sum_{k=1}^{K}\hat{f}_{-k}(\mathbf{X}_i)\hat{f}_{-k}(\mathbf{X}_j)-\frac{1}{K}\sum_{k=1}^{K}\hat{f}_{-k}(\mathbf{X}_i)\frac{1}{K}\sum_{k=1}^{K}\hat{f}_{-k}(\mathbf{X}_j).$$

Thus, we have the following cross-validation estimator of the required term,

$$\begin{aligned}&\mathbb{E}\left\{\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i),\widehat{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)}\mid\mathbf{X}_i,\mathbf{X}_j)Y_i(s)Y_j(t)\right\}\\ =\;&\frac{\sum_{i=1}^{n}\sum_{j\neq i}\mathbf{1}\{T_i=s,T_j=t\}Y_iY_j\cdot\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i),\widehat{\hat{f}_{\mathcal{Z}^{tr}}}(\mathbf{X}_j)\mid\mathbf{X}_i,\mathbf{X}_j)}{\sum_{i=1}^{n}\sum_{j\neq i}\mathbf{1}\{T_i=s,T_j=t\}}.\end{aligned}$$

for $s,t\in\{0,1\}$. However, since this naive calculation is computationally expensive, we rewrite it as follows to reduce the computational time from $O(n^2K)$ to $O(nK)$,

$$\frac{\sum_{k=1}^{K}\left(\sum_{i=1}^{n}\mathbf{1}\{T_i=s\}Y_i\hat{f}_{-k}(\mathbf{X}_i)\right)\left(\sum_{i=1}^{n}\mathbf{1}\{T_i=t\}Y_i\hat{f}_{-k}(\mathbf{X}_i)\right)-\sum_{i=1}^{n}\mathbf{1}\{T_i=s,T_i=t\}Y_i^2\hat{f}_{-k}(\mathbf{X}_i)}{K\left[\left(\sum_{i=1}^{n}\mathbf{1}\{T_i=s\}\right)\left(\sum_{i=1}^{n}\mathbf{1}\{T_i=t\}\right)-\sum_{i=1}^{n}\mathbf{1}\{T_i=s,T_i=t\}\right]}$$

$$- \frac{(\sum_{i=1}^n \sum_{k=1}^K \mathbf{1}\{T_i = s\} Y_i \hat{f}_{-k}(\mathbf{X}_i))(\sum_{i=1}^n \sum_{k=1}^K \mathbf{1}\{T_i = t\} Y_i \hat{f}_{-k}(\mathbf{X}_i))}{K^2 \left[(\sum_{i=1}^n \mathbf{1}\{T_i = s\})(\sum_{i=1}^n \mathbf{1}\{T_i = t\}) - \sum_{i=1}^n \mathbf{1}\{T_i = s, T_i = t\}\right]}$$
$$+ \frac{\sum_{i=1}^n \sum_{k=1}^K \mathbf{1}\{T_i = s, T_i = t\} Y_i^2 \hat{f}_{-k}(\mathbf{X}_i)}{K^2 \left[(\sum_{i=1}^n \mathbf{1}\{T_i = s\})(\sum_{i=1}^n \mathbf{1}\{T_i = t\}) - \sum_{i=1}^n \mathbf{1}\{T_i = s, T_i = t\}\right]}.$$

### A.5.2 The Population Average Prescriptive Effect (PAPE)

Next, we propose the following cross-validation estimator of the PAPE for an estimated ITR (equation (A10)),

$$\hat{\tau}_F = \frac{1}{K} \sum_{k=1}^K \hat{\tau}_{\hat{f}_{-k}}(\mathcal{Z}_k) \qquad \text{(A13)}$$

where $\hat{\tau}_f(\cdot)$ is defined in equation (A2). The next theorem shows the unbiasedness of this estimator and derives its variance.

THEOREM A5 (UNBIASEDNESS AND EXACT VARIANCE OF THE CROSS-VALIDATION PAPE ESTIMATOR) *Under Assumptions 1, 2, and 3, the expectation and variance of the cross-validation PAPE estimator defined in equation* (A13) *are given by,*

$$\begin{aligned}
\mathbb{E}(\hat{\tau}_F) &= \tau_F \\
\mathbb{V}(\hat{\tau}_F) &= \frac{m^2}{(m-1)^2}\Bigg[\frac{\mathbb{E}(\widetilde{S}_{\hat{f}1}^2)}{m_1} + \frac{\mathbb{E}(\widetilde{S}_{\hat{f}0}^2)}{m_0} + \frac{1}{m^2}\left\{\tau_F^2 - m p_F(1-p_F)\tau^2 + 2(m-1)(2p_F-1)\tau\tau_F\right\} \\
&\quad + \frac{1}{m^2}\mathbb{E}\Big\{\left\{(m-3)(m-2)\tau^2 + (m^2-2m+2)\tau_i\tau_j - 2(m-2)^2\tau\tau_i\right\} \\
&\quad \times \mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\Big\}\Bigg] - \frac{K-1}{K}\mathbb{E}(\widetilde{S}_F^2)
\end{aligned}$$

*for* $i \neq j$*, where* $\widetilde{S}_{\hat{f}t}^2 = \sum_{i=1}^m (\widetilde{Y}_{\hat{f}i}(t) - \overline{\widetilde{Y}_{\hat{f}}(t)})^2/(m-1)$*,* $\widetilde{S}_F^2 = \sum_{k=1}^K (\hat{\tau}_{\hat{f}_{-k}}(\mathcal{Z}_k) - \overline{\hat{\tau}_{\hat{f}_{-k}}(\mathcal{Z}_k)})^2/(K-1)$ *with* $\widetilde{Y}_{\hat{f}i}(t) = (\hat{f}_{-k}(\mathbf{X}_i) - \hat{p}_{\hat{f}_{-k}})Y_i(t)$*,* $\overline{\widetilde{Y}_{\hat{f}}(t)} = \sum_{i=1}^m \widetilde{Y}_{\hat{f}i}(t)/m$*, and* $\overline{\hat{\tau}_{\hat{f}_{-k}}(\mathcal{Z}_k)} = \sum_{k=1}^K \hat{\tau}_{\hat{f}_{-k}}(\mathcal{Z}_k)/K$*, for* $t = \{0, 1\}$*.*

**Proof** The proof of unbiasedness is similar to that of Appendix A.1.2 and thus is omitted. The derivation of the variance is similar to that of Appendix A.5.1. The key difference is the calculation of the following covariance term, which needs care due to the randomness of $\hat{f}_{\mathcal{Z}^{tr}}$,

$$\begin{aligned}
&\mathrm{Cov}(Y_i^*(1) - Y_i^*(0), Y_j^*(1) - Y_j^*(0)) \\
&= \mathrm{Cov}\left\{\left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) - \frac{1}{m}\sum_{i'=1}^m \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_{i'})\right)\tau_i, \left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) - \frac{1}{m}\sum_{j'=1}^m \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_{j'})\right)\tau_j\right\},
\end{aligned}$$

where $i \neq j$. There are seven terms that need to be carefully expanded,

$$\begin{aligned}
&\frac{m-2}{m^2}\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_k)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_k)\tau_j) + \frac{(m-2)(m-3)}{m^2}\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_k)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_\ell)\tau_j) \\
&+ \frac{(m-1)^2}{m^2}\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\tau_j) + \frac{1}{m^2}\mathrm{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_j)
\end{aligned}$$

$$
\begin{aligned}
&- \frac{2(m-1)}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_j) + \frac{2(m-2)}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_k)\tau_j) \\
&- \frac{2(m-2)(m-1)}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_k)\tau_j),
\end{aligned}
$$

where $i, j, k, \ell$ represent indices that do not take an identical value at the same time (e.g., $i \neq j$). Then, we rewrite the above terms using the properties of covariance as follows,

$$
\begin{aligned}
& \frac{(m-2)\tau^2}{m^2}\mathbb{V}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)) + \frac{(m-2)(m-3)\tau^2}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)) \\
&+ \frac{(m-1)^2}{m^2}\left[\mathbb{E}\left\{\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right\} + \operatorname{Cov}(\bar{f}_F(\mathbf{X}_i)\tau_i, \bar{f}_F(\mathbf{X}_j)\tau_j)\right] \\
&+ \frac{1}{m^2}\left[\mathbb{E}\left\{\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right\} + \operatorname{Cov}(\bar{f}_F(\mathbf{X}_j)\tau_i, \bar{f}_F(\mathbf{X}_i)\tau_j)\right] \\
&- \frac{2(m-1)\tau}{m^2}(1-p_F)\mathbb{E}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)\tau_i) \\
&+ \frac{2(m-2)\tau}{m^2}\left[\mathbb{E}\left\{\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i] + p_F \operatorname{Cov}(\bar{f}_F(\mathbf{X}_i), \tau_i)\right\}\right] \\
&- \frac{2(m-2)(m-1)\tau}{m^2}\left[\mathbb{E}\left\{\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\right\} + \operatorname{Cov}(\bar{f}_F(\mathbf{X}_i)\tau_i, \bar{f}_F(\mathbf{X}_j))\right] \\
=\; & \frac{(m-2)\tau^2}{m^2}p_F(1-p_F) + \frac{(m-2)(m-3)\tau^2}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j)) \\
&+ \frac{m^2-2m+2}{m^2}\mathbb{E}\left\{\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right\} + \frac{1}{m^2}(\tau_F^2 + 2\tau p_F \tau_F) \\
&- \frac{2(m-1)\tau^2}{m^2}p_F(1-p_F) - \frac{2(m-1)\tau\tau_F}{m^2}(1-p_F) + \frac{2(m-2)\tau}{m^2}p_F\tau_F \\
&- \frac{2(m-2)^2\tau}{m^2}\mathbb{E}\left\{\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_k) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\right\} \\
=\; & \frac{1}{m^2}\left(\tau_F^2 - mp_F(1-p_F)\tau^2 + 2(m-1)(2p_F-1)\tau\tau_F\right) \\
&+ \mathbb{E}\left[\frac{(m-2)(m-3)\tau^2}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\right. \\
&\qquad - \frac{2(m-2)^2\tau}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i. \\
&\qquad\quad \left. + \frac{m^2-2m+2}{m^2}\operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j) \mid \mathbf{X}_i, \mathbf{X}_j)\tau_i\tau_j\right],
\end{aligned}
$$

for $i \neq j$, where we used the results from Appendix A.1.2 as well as $\mathbb{V}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i)) = p_F(1-p_F)$ and $\tau_F = \operatorname{Cov}(\bar{f}_F(\mathbf{X}_i), \tau_i) = \operatorname{Cov}(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i), \tau_i)$. □

Like the case of the population average value, the variance has two extra terms when compared to the case of a fixed ITR (see Theorem A2 in Appendix A.1.2). The estimation of the variance is similar to that for the population average value.

## A.6 Proof of Theorem 3

We begin by deriving the variance. The derivation proceeds in the same fashion as the one for Theorem A5 (see Appendix A.5.2). The only non-trivial change is the derivation of the covariance term, which we detail below. First, similar to the proof of Theorem 1 (see

Appendix A.2), we note the following relation:

$$\begin{aligned}
\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p) = 1) &= \mathbb{E}\left(\int_{-\infty}^{\infty} \Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c) = 1 \mid \mathcal{Z}^{tr}, \hat{c}_p = c) P(\hat{c}_p = c \mid \mathcal{Z}^{tr}) \mathrm{d}c\right) \\
&= \mathbb{E}\left(\int_{-\infty}^{\infty} \frac{\lfloor mp \rfloor}{m} P(\hat{c}_p = c \mid \mathcal{Z}^{tr}) \mathrm{d}c\right) \\
&= \frac{\lfloor mp \rfloor}{m},
\end{aligned}$$

where the second equality follows from the fact that conditioned on a fixed training set $\mathcal{Z}^{tr}$ and conditioned on $\hat{c}_p = c$, exactly $\lfloor mp \rfloor$ out of $m$ units will be assigned to the treatment condition. Given this result, we can compute the covariance as follows,

$$\begin{aligned}
& \operatorname{Cov}(\widetilde{Y}_i(1) - \widetilde{Y}_i(0), \widetilde{Y}_j(1) - \widetilde{Y}_j(0)) \\
= \ & \operatorname{Cov}\left\{\left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p) - p\right)\tau_i, \left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j, \hat{c}_p) - p\right)\tau_j\right\} \\
= \ & \operatorname{Cov}\left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p)\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j, \hat{c}_p)\tau_j\right) - 2p\operatorname{Cov}\left(\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j, \hat{c}_p)\tau_j\right) \\
= \ & \frac{m\lfloor mp \rfloor(\lfloor mp \rfloor - 1) - \lfloor mp \rfloor^2(m-1)}{m^2(m-1)} \mathbb{E}(\tau_i \mid \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p) = 1)^2 - 2p\operatorname{Cov}\left(\tau_i, \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_j, \hat{c}_p)\tau_j\right) \\
= \ & \frac{\lfloor mp \rfloor(\lfloor mp \rfloor - m)}{m^2(m-1)}\kappa_{F1}(p)^2 + \frac{2p\lfloor mp \rfloor(m - \lfloor mp \rfloor)}{m^2(m-1)}\left(\kappa_{F1}(p)^2 - \kappa_{F1}(p)\kappa_{F0}(p)\right) \\
= \ & (2p-1)\frac{\lfloor mp \rfloor(m - \lfloor mp \rfloor)}{m^2(m-1)}\kappa_{F1}(p)^2 - \frac{2p\lfloor mp \rfloor(m - \lfloor mp \rfloor)}{m^2(m-1)}\kappa_{F1}(p)\kappa_{F0}(p) \\
= \ & \frac{\lfloor mp \rfloor(m - \lfloor mp \rfloor)}{m^2(m-1)}\left\{(2p-1)\kappa_{F1}(p)^2 - 2p\kappa_1(p)\kappa_{F0}(p)\right\}
\end{aligned}$$

Combining this covariance result with the expression for the marginal variances yields the desired variance expression for $\hat{\tau}_{Fp}$.

Next, we derive the upper bound of bias. Using the same technique as the proof of Theorem A2, we can rewrite the expectation of the proposed estimator as,

$$\mathbb{E}(\hat{\tau}_{Fp}) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p)\right) - pY_i(1) - (1-p)Y_i(0)\right\}\right]$$

Now, define $F(c) = \mathbb{P}(s(\mathbf{X}_i) \leq c)$. Without loss of generality, assume $\hat{c}_p > c_p$ (If this is not the case, we simply switch the upper and lower limits of the integrals below). Then, the bias of the estimator is given by,

$$\begin{aligned}
|\mathbb{E}(\hat{\tau}_{Fp}) - \tau_{Fp}| &= \left|\mathbb{E}\left\{\mathbb{E}(\hat{\tau}_{Fp} - \tau_{Fp} \mid \mathcal{Z}^{tr})\right\}\right| \\
&\leq \mathbb{E}\left\{\left|\mathbb{E}(\hat{\tau}_{Fp} - \tau_{Fp} \mid \mathcal{Z}^{tr})\right|\right\} \\
&= \mathbb{E}\left\{\left|\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p)\right) - Y_i\left(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, c_p)\right)\right\} \;\middle|\; \mathcal{Z}^{tr}\right]\right|\right\} \\
&= \mathbb{E}\left\{\left|\mathbb{E}\left[\int_{c_p}^{\hat{c}_p} \mathbb{E}(\tau_i \mid \hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = c, \mathcal{Z}^{tr}) \mathrm{d}F(c)\right]\right|\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left\{\left|\mathbb{E}\left[\int_{F(c_p)}^{F(\hat{c}_p)} \mathbb{E}(\tau_i \mid \hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = F^{-1}(x), \mathcal{Z}^{tr})\mathrm{d}x\right]\right|\right\} \\
&\leq \mathbb{E}\left[|F(\hat{c}_p) - 1 - p| \times \max_{c\in[c_p,\hat{c}_p]} \left|\mathbb{E}(\tau_i \mid \hat{s}_{\mathcal{Z}^{tr}}(\mathbf{X}_i) = c, \mathcal{Z}^{tr})\right|\right].
\end{aligned}
$$

By the definition of $\hat{c}_p$, $F(\hat{c}_p)$ is the $m - \lfloor mp \rfloor$th order statistic of $n$ independent uniform random variables. This statistic does not depend on the training set $\mathcal{Z}^{tr}$ as the test samples are independent. Thus, $F(\hat{c}_p)$ follows the Beta distribution with the shape and scale parameters equal to $m - \lfloor mp \rfloor$ and $\lfloor mp \rfloor + 1$, respectively. Therefore, we have,

$$
\mathbb{P}(|F(\hat{c}_p) - p| > \epsilon) = 1 - B(1-p+\epsilon, m-\lfloor mp \rfloor, \lfloor mp \rfloor+1) + B(1-p-\epsilon, m-\lfloor mp \rfloor, \lfloor mp \rfloor+1), \tag{A14}
$$

where $B(\epsilon, \alpha, \beta)$ is the incomplete beta function, i.e.,

$$
B(\epsilon, \alpha, \beta) = \int_0^{\epsilon} t^{\alpha-1}(1-t)^{\beta-1}\mathrm{d}t.
$$

Combining with the result above, the desired result follows. □

## A.7 The Population Average Prescriptive Difference of Estimated ITRs under a Budget Constraint

We consider the estimation and inference for the PAPD of an estimated ITR. The cross-validation estimator of this quantity is given by,

$$
\widehat{\Delta}_p(F, G) = \frac{1}{K}\sum_{k=1}^{K} \widehat{\Delta}_p(\hat{f}_{-k}, \hat{g}_{-k}, \mathcal{Z}_k), \tag{A15}
$$

where $\widehat{\Delta}_p(\hat{f}_{-k}, \hat{g}_{-k}, \mathcal{Z}_k)$ is defined in equation (8). Although the bias of the proposed estimator is not zero, we derive its upper bound as done in Theorem 3.

THEOREM A6 (BIAS AND VARIANCE OF THE CROSS-VALIDATION PAPD ESTIMATOR WITH A BUDGET CONSTRAINT) *Under Assumptions 1, 2, and 3, the bias of the cross-validation PAPD estimator with a budget constraint $p$ defined in equation* (A15) *can be bounded as follows,*

$$
\begin{aligned}
&\mathbb{E}_{\mathcal{Z}^{tr}}[\mathbb{P}_{c_p(\hat{f}_{\mathcal{Z}^{tr}}), c_p(\hat{g}_{\mathcal{Z}^{tr}})}(|\mathbb{E}\{\widehat{\Delta}_p(F,G) - \Delta_p(F,G) \mid c_p(\hat{f}_{\mathcal{Z}^{tr}}), c_p(\hat{g}_{\mathcal{Z}^{tr}})\}| \geq \epsilon)] \\
&\leq 1 - 2B(1-p+\gamma_p(\epsilon), n-\lfloor np \rfloor, \lfloor np \rfloor+1) + 2B(1-p-\gamma_p(\epsilon), n-\lfloor np \rfloor, \lfloor np \rfloor+1),
\end{aligned}
$$

*where any given constant $\epsilon > 0$, $B(\epsilon, \alpha, \beta)$ is the incomplete beta function (if $\alpha = 0$ and $\beta > 0$, we set $B(\epsilon, \alpha, \beta) := H(\epsilon)$ for all $\epsilon$ where $H(\epsilon)$ is the Heaviside step function), and*

$$
\gamma_p(\epsilon) = \frac{\epsilon}{\mathbb{E}_{\mathcal{Z}^{tr}}[\max_{c\in[c_p(\hat{f}_{\mathcal{Z}^{tr}})-\epsilon, c_p(\hat{f}_{\mathcal{Z}^{tr}})+\epsilon], d\in[c_p(\hat{g}_{\mathcal{Z}^{tr}})-\epsilon, c_g(\hat{g}_{\mathcal{Z}^{tr}})+\epsilon]}\{\mathbb{E}(\tau_i \mid \hat{s}_{\hat{f}_{\mathcal{Z}^{tr}}}(\mathbf{X}_i) = c), \mathbb{E}(\tau_i \mid \hat{s}_{\hat{g}_{\mathcal{Z}^{tr}}}(\mathbf{X}_i) = d)\}]}.
$$

*The variance of the estimator is,*

$$
\mathbb{V}(\widehat{\Delta}_p(F,G)) = \frac{\mathbb{E}(S^2_{\hat{f}\hat{g}1})}{n_1} + \frac{\mathbb{E}(S^2_{\hat{f}\hat{g}0})}{n_0} + \frac{\lfloor np \rfloor(\lfloor np \rfloor - n)}{n^2(n-1)}(\kappa_{F1}(p)^2 + \kappa_{G1}(p)^2)
$$

$$
-2\left(\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = \hat{g}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = 1) - \frac{\lfloor np \rfloor^2}{n^2}\right)\kappa_{F1}(p)\kappa_{G1}(p)
$$
$$
-\frac{K-1}{K}\mathbb{E}(S_{FG}^2),
$$

*where* $S_{\hat{f}\hat{g}t}^2 = \sum_{i=1}^n (\widetilde{Y}_{\hat{f}\hat{g}i}(t) - \overline{\widetilde{Y}_{\hat{f}\hat{g}}(t)})^2/(n-1)$, $S_{FG}^2 = \sum_{k=1}^K (\widehat{\Delta}_p(\hat{f}_{-k}, \hat{g}_{-k}, \mathcal{Z}_k)) - \overline{\widehat{\Delta}_p(\hat{f}_{-k}, \hat{g}_{-k}, \mathcal{Z}_k)})^2/(K-1)$, $\kappa_{Ft}(p) = \mathbb{E}(\tau_i \mid \hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = t)$*, and* $\kappa_{Gt}(p) = \mathbb{E}(\tau_i \mid \hat{g}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(g)) = t)$ *with* $Y_{\hat{f}\hat{g}i}(t) = \left\{\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) - \hat{g}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{g}_{\mathcal{Z}^{tr}}))\right\} Y_i(t)$, $\overline{Y(t)} = \sum_{i=1}^n Y_{\hat{f}\hat{g}i}(t)/n$*, and* $\overline{\widehat{\Delta}_p(\hat{f}_{-k}, \hat{g}_{-k}, \mathcal{Z}_k)} = \sum_{k=1}^K \widehat{\Delta}_p(\hat{f}_{-k}, \hat{g}_{-k}, \mathcal{Z}_k)/K$ *for* $t = 0, 1$.

Proof is similar to that of Theorem 3, and hence is omitted. To estimate the variance, it is tempting to replace all unknowns with their sample analogues. However, the empirical analogue for the joint probability $\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = \hat{g}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{g}_{\mathcal{Z}^{tr}})) = 1)$ under general $\hat{f}_{\mathcal{Z}^{tr}}, \hat{g}_{\mathcal{Z}^{tr}}$ is not a good estimate because it is solely based on one realization. Thus, we use the following conservative bound,

$$
\begin{aligned}
&-\left(\Pr(\hat{f}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = \hat{g}_{\mathcal{Z}^{tr}}(\mathbf{X}_i, \hat{c}_p(\hat{f}_{\mathcal{Z}^{tr}})) = 1) - \frac{\lfloor np \rfloor^2}{n^2}\right)\kappa_{F1}(p)\kappa_{G1}(p) \\
&\leq \frac{\lfloor np \rfloor \max\{\lfloor np \rfloor, n - \lfloor np \rfloor\}}{n^2(n-1)}|\kappa_{F1}(p)\kappa_{G1}(p)|,
\end{aligned}
$$

where the inequality follows because the maximum is achieved when the scoring rules of $\hat{f}_{\mathcal{Z}^{tr}}$ and $\hat{g}_{\mathcal{Z}^{tr}}$ are perfectly negatively correlated. We use this upper bound in our simulation and empirical studies. In Section 5, we find that this upper bound estimate of the variance produces only a small conservative bias.

## A.8 An Additional Empirical Application

In this section, we describe an additional empirical application based on the canvassing experiment (Broockman and Kalla, 2016). This study was also re-analyzed by Künzel *et al.* (2018). The original authors find little heterogeneity in treatment effect. Our analysis below confirms this finding.

### A.8.1 The Experiment and Setup

We analyze the transgender canvas study of Broockman and Kalla (2016). This is an experiment that randomly assigned a door-to-door canvassing treatment to over 1,200 households (with a total of over 1,800 members) in Florida to estimate the treatment effect on support for a transgender rights law. The placebo group received a conversation on recycling, while the treatment group received a conversation about transgender issues. The support is measured at various time points after the intervention (i.e., 3 days, 3 weeks, 6 weeks, 3 months) using an online survey. The treatment effect heterogeneity is important in this scenario as canvassing is both costly and time-consuming. An ITR may allow canvassers to contact only those who are positively influenced by the message.

We follow the pre-experiment analysis plan by the original authors, and select a total of 26 baseline covariates including political inclination, gender, race, and opinions on social

| | BCF | | | Causal Forest | | | R-Learner | | |
|---|---|---|---|---|---|---|---|---|---|
| | est. | s.e. | treated | est. | s.e. | treated | est. | s.e. | treated |
| No budget constraint | $-0.104$ | $0.128$ | $48.4\%$ | $-0.349$ | $0.137$ | $47.5\%$ | $0$ | $0$ | $100\%$ |
| $20\%$ Budget constraint | $-0.02$ | $0.121$ | $20\%$ | $-0.120$ | $0.107$ | $20\%$ | $0$ | $0.104$ | $20\%$ |

Table A2: The Estimated Population Average Prescription Effect (PAPE) for Bayesian Causal Forest (BCF), Causal Forest, and R-Learner with and without a Budget Constraint. For each of the three outcomes, the point estimate, the standard error, and the proportion treated are shown. The budget constraint considered here implies that the maximum proportion treated is 20%.

issues. Our treatment variable is whether or not the individual received the conversation about transgender issues (as opposed to the recycling message). Since the randomization was conducted on the household level, we randomly select one individual from each household for our analysis. We focus on the primary target (support for the transgender law) at the 3 day time point after the intervention, which is measured on a discrete scale with 7 possible values $\{-3, -2, -1, 0, 1, 2, 3\}$, with positive values indicating support.

The resulting dataset consists of 409 observations. We randomly select approximately 70% of the sample (i.e., 287 observations) as the training data and the reminder of the sample (i.e., 122 observations) as the evaluation data. We center the outcome variable using the mean in the training data to minimize variance, as discussed in Section 2. We train three machine learning models designed to measure heterogeneous treatment effects: Causal Forests, Bayesian Causal Forests (Hahn *et al.*, 2020), and R-Learner (Nie and Wager, 2017). All tuning was done through the 5-fold cross validation procedure on the training set using the PAPE as the evaluation metric. For Causal Forest, we set `tune.parameters = TRUE`. For Bayesian Causal Forests (BCF), tuning was done using a burn-in sample for MCMC sampling. For R-Learner, we utilized the lasso loss function and the default cross-validation for the regularization parameter. We then create an ITR as $\mathbf{1}\{\hat{\tau}(\mathbf{x}) > 0\}$ where $\hat{\tau}(\mathbf{x})$ is the estimated conditional average treatment effect obtained from each fitted model. We will evaluate these ITRs using the evaluation sample.

### A.8.2 Results

Table A2 presents the results. We find that without a budget constraint, none of the ITRs based on the machine learning methods significantly improves upon the random treatment rule. In particular, the R-Learner leads to an ITR that treats everyone. Furthermore, we see that the ITR based on Causal Forest performs worse than the random treatment rule by 0.349 (out of a $-3$ to 3 scale) with a standard error of 0.137. The results are similar when we impose a budget constraint, and none of the resulting ITRs perform statistically significantly better than the random treatment rule. The result based on R-Learner is consistent with the conclusion of the original study indicating that there was no heterogeneity detected using LASSO.

We plot the estimated PAPE (with 95% pointwise confidence interval) as a function of budget constraint in Figure A1. The area between this line and the black horizontal line at zero corresponds to the AUPEC. In each plot, the horizontal axis represents the budget

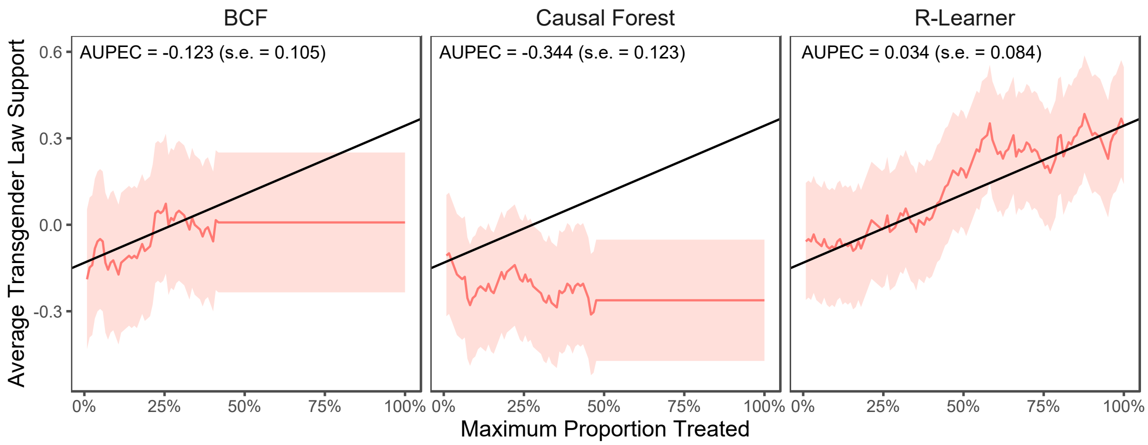


Figure A1: Estimated Area Under the Prescriptive Effect Curve (AUPEC). A solid red line in each plot represents the Population Average Prescriptive Effect (PAPE) with pointwise 95% confidence intervals shaded. The area between this line and the black line (representing random treatment) is the AUPEC. The results are presented for the individualized treatment rules based on Bayesian Causal Forest (BCF), Causal Forest, and R-Learner.

constraint as the maximum proportion treated, and the point estimate and standard error of the AUPEC are shown. While BCF and R-Learner fail to create an ITR that is significantly different from the random treatment rule, Causal Forest produces an ITR that is statistically significantly worse than the random treatment rule. The result illustrates a potential danger of using an advanced machine learning algorithm to create an ITR. Indeed, there is no guarantee that the resulting ITR outperforms the random treatment rule.