Policy Learning with Compliance Guarantee

Thomas Chan, Vadim Marmer and Kyungchul Song
Vancouver School of Economics, University of British Columbia

Preliminary Draft: Do Not Distribute without Permission.

Abstract. We study optimal policy learning where a policy maker uses policy outcome data from a source population to design treatment assignments for a target population under budget constraint. Due to the budget constraint, the policy maker needs to consider both the treatment effects and individuals' incentives for treatment participation to minimize wasted resources. The main challenge is that treatment participation incentives may differ between the two populations. We develop a maximin approach that maximizes the minimum expected treatment outcome across all possible incentive configurations. We find that this optimal policy learning problem transforms into one with stochastic dominance constraints, where optimal assignment prioritizes individuals most likely to comply with the treatment assignment.

Key words. Policy Learning; Stochastic Dominance; Compliance Guarantee

JEL Classification:

Date: June 30, 2025.
We are grateful to Gordon Anderson and Sukjin Han and participants in Stochastic Dominance and Quantile-based Methods in Financial Econometrics Workshop and the CUHK-Shenzhen Econometrics Workshop for valuable comments and suggestions. Song acknowledges the support by Social Sciences and Humanities Research Council of Canada.

1. Introduction

Policy learning refers to designing an optimal policy for a target population using data from a source population. As such, it is inherently a *transfer problem*, which involves transferring features learned about one population to another. For example, suppose that a policymaker (PM) has access to experimental data from a source population. From the experiment data, PM can estimate the treatment effect for the source population, and wants to obtain an optimal treatment assignment for a target population. Of course, the relevance of the source population data to the treatment assignment problem in a different population depends on how much of the experiment setting is "transferable" to the target population. One extreme setting, often considered in the literature, is that both populations represent the same population, regarding the data as generated from the target population. However, such an assumption is not plausible in actual policy settings.

It is not uncommon in the literature of randomized experiments that the causal effects of a certain program estimated to be statistically significant disappear when the program is expanded to a wider population with an increased budget (Duflo (2004), Allcott (2015), Muralidharan and Niehaus (2017), Wang and Yang (2021)). In summary, the experiment results can lose their relevance for a PM targeting a wider population, if the source and target populations are very different.

A departure from the single-population assumption, acknowledging the issue of transfer from the source to target population, opens up a wide range of possibilities. In this paper, we consider a policy learning setting under budget constraints. In this setting, we assume transferability between target and source populations, where the conditional distributions of potential outcomes and total costs given covariates are transferable between the source and target populations but the program participation incentives are not. In other words, the source and target populations may exhibit different take-up behaviors. This could be, for example, because the experiments are administered by the NGO workers who have high stakes at the success of the program and put extra efforts to raise the take-up rate, whereas the program is administered by the civil servants in the government with less stakes in the program's success. In this case, the take-up behaviors between the two populations can be widely different.

In our setting, the PM is subject to a budget constraint. The size of the budget depends on the take-up rate of the subjects. Hence, the optimal decision would take this into account and consider both the take-up rate and the contribution of treatment to the social objective function. However, we consider a setting where the policy has not been implemented in the population. In this setting, it is too strong an assumption in practice that the PM knows the participation incentives on the target population. Thus, we relax this assumption, allowing the

PM to be ignorant of the actual take-up rate in the population. In formulating the policy design problem, we take a maximin approach where the PM aims to find a policy that maximizes the minimum social objective across a range of participation incentives.

First, we specify that the participation decisions are made through a generalized Roy model where each subject decides to participate in the program based on the net expected payoff differences. However, the PM does not know the individual payoff functions, except that they belong to a certain class of utility functions. The size of the class is regulated through a bound on the maximal variations in the marginal utilities. Then, we find that the optimal policy design problem in such a setting can be reformulated as one under the almost first order stochastic dominance constraint of Leshno and Levy (2002). From this reformulation, we obtain an explicit form of an optimal policy. This policy is intuitive: it focuses on those that are expected to fully comply and prioritize those with highest returns to the treatment that is defined to be the ratio of the conditional treatment effect to the conditional total costs given covariates. Then, we turn to an estimator of the optimal policy using the combined sample of the source population and the target population, and establishes its consistency.

The optimal policy design has received attention in the literature. Manski (2004) proposed an optimal policy that maximizes the empirical welfare and suggested using maximal regret as its performance measure. Kitagawa and Tetenov (2018) established finite sample upper and lower bounds for the finite sample maximal regret for the empirical welfare maximizer. Athey and Wager (2021) considered a setting where the source population data are generated from an observational setting, and proposed an optimal policy motivated from methods of semiparametric efficiency. A recent work by Chernozhukov, Lee, Rosen, and Sun (2025) proposes an upper confidence bound (UCB) approach in policy learning that explicitly incorporates the estimation error as part of the policy objective function. This literature incorporates a constrained optimization problem by considering the restricted set of policy functions.[1]

There is a line of research which explicitly considers budget constraints. Bhattacharya and Dupas (2012) considered a problem of welfare maximizing treatment assignment and derived an optimal solution under budget constraints. More recently, Sun, Munro, Kalashnov, Du, and Wager (2025) considered a setting where the PM does not know exactly the costs. Sun (2025) studied an empirical welfare maximization problem under budget constraints. The policy learning problem under the budget constraint is mathematically similar to that under fairness considerations (Viviano and Bradic (2024)). Our paper is distinct from this literature, by allowing the target population to have different participation incentives from the source population.

[1]There is related literature on external validity on randomized experiment results. See, e.g., Hotz, Imbens, and Mortimer (2005),Gechter (2024), and Gechter and Meager (2022).

The paper is organized as follows. In the next section, we present the policy learning problem, making explicit the transferability condition. Toward the optimal policy solution, we first provide a reformulation of the policy design problem using almost first order stochastic dominance constraints, and then present an optimal policy at the population level. In this section, we also propose an estimated optimal policy using the combined sample from the target and source populations, and establish its consistency. We provide numerical illustration of the optimal policy using simulations. In Section 3, we discuss some extensions. In Section 4, we conclude. The mathematical proofs are found in the appendix.

2. Optimal Treatment Assignments with Compliance Guarantee

2.1. The Policy Objectives

In this section, we introduce the basic policy learning set-up. In the target population, each individual is endowed with the potential outcomes $Y(1)$ and $Y(0)$, where $Y(1)$ indicates the potential outcome of the treated state and $Y(0)$ that of the control state. To formally model the incentives, we assume that each individual is endowed with the utility $u(y,v)$, when the potential outcome of the treatment and the other payoff states are realized to be y and v. Each individual observes the payoff state V as a random vector, but does not observe the potential outcome of the treatment at the time of deciding on the participation in the treatment program. Throughout the paper, we assume that the payoff state V taking values in a space $\mathcal{V}$ includes a covariate vector X so that

$$V = (X, V_{-X}),$$

where V_{-X} denotes the payoff states other than X. We assume that the PM observes X but does not observe V_{-X}. Thus, the random utility of an agent with the outcome y is given by

$$u(y, V).$$

When offered treatment, an individual decides whether to accept it or not by comparing the expected utilities between the treatment state and control state. Thus, the acceptance decision can be written as

$$a(V;u) := \mathbf{1}\{\mathbf{E}[u(Y(1),V) \mid V] \geq \mathbf{E}[u(Y(0),V) \mid V]\},$$

where the conditional expectation given V reflects that the individual observes V, knows their own utility function u, and has rational expectations.

A policy by the PM is a map: $g : \mathcal{X} \to \{0,1\}$, which assigns each covariate group $X = x$ to the treatment group ($g(x) = 1$) or the control group ($g(x) = 0$). In considering the outcome, we assume that the PM adopts a *social objective function*, $q(y,x):\mathcal{Y}\times\mathcal{X}\to\mathbf{R}$, so that for each realized outcome, the PM deems the outcome y more desirable for group x if $q(y,x)$ is higher. We will discuss some examples of q later. Thus, under the policy g, the PM focuses on the following aggregate quantity:

$$S(g;u):=\mathbf{E}\Big[q(Y(0),X)\Big(1-g(X)a(V;u)\Big)+q(Y(1),X)g(X)a(V;u)\Big]. \tag{2.1}$$

The expected social outcome not only depends on the policy g but also the take-up decision a by individuals.

As for the choice of q, we provide two examples as follows:

(Weighted Average): $q(y,x)=q_W(y):=y\cdot w(x)$.
(Sufficientarianism): $q(y,x)=q_S(y):=\min\{y,y^*\}\cdot w(x)$.

The choice $q(y)=q_W(y)$ says that the PM considers the treatment a success if the weighted average of the outcomes for those treated is high. On the other hand, by choosing $q(y)=q_S(y)$, the PM takes as a measure of success the average outcome below a threshold y^*. This measure reflects sufficientarianism which suggests that the PM should not be concerned about the outcome ordering among those groups with high enough outcomes (see Alcantud, Mariotti, and Veneziani (2022) and Bossert, Cato, and Kamaga (2023)).[2]

The policy is subject to a budget constraint. We consider two kinds of costs, C_O and C_T, where the random variable C_O represents the cost incurred by sending out an offer of treatment and C_T the cost incurred by treating the individual. The budget constraint is written as:

$$\mathbf{E}[C_O g(X)]+\mathbf{E}[C_T g(X)a(V;u)]\le B, \tag{2.2}$$

where B denotes the budget allowed for the program. Thus, the main goal of PM in the target population is to maximize the expected outcome in (2.1) subject to the budget constraint (2.2).

However, the PM cannot implement this policy learning problem (even at the population level), primarily because the PM does not know the participation incentives of the target population. We consider a maximin approach, where the PM maximizes the expected outcome for the least favorable group with participation incentives that minimizes the expected outcome.

Let us formalize the PM's problem at the population level. Define

$$Q(g;u)=\begin{cases} S(g;u), & \text{if (2.2) is satisfied,}\\ -\infty, & \text{otherwise.}\end{cases} \tag{2.3}$$

[2]We are grateful to Gordon Anderson for introducing the notion of sufficientarianism and the relevant literature to us.

Then, the PM considers the optimal policy that maximizes the minimum welfare over $g \in \mathcal{G}$ subject to the budget constraint:

$$\sup_{g\in\mathcal{G}} \inf_{u\in\mathcal{U}} Q(g;u), \tag{2.4}$$

where $\mathcal{U}$ denotes the class of utility functions that we explain in more detail later. Here, we consider $\mathcal{G}$ to be the collection of (measurable) maps from $\mathcal{X}$ to $\{0,1\}$. Then, a policy g^* is optimal if it satisfies the following:

$$\inf_{u\in\mathcal{U}} Q(g^*;u) = \sup_{g\in\mathcal{G}} \inf_{u\in\mathcal{U}} Q(g;u). \tag{2.5}$$

The problem appears complex, because u is involved both in the objective function and the budget constraint. Later, we will show how we can reformulate the problem into a tractable one.

It is worth noting that the PM does not consider the random utility $u(y,V)$ as part of the PM's objective function, except in the participation decisions by the individuals. We have several reasons to formulate the PM's decision problem this way. First, it may not be obvious for the PM to come up with a social utility function out of the subjective utilities in a way that is transparent and socially agreeable. Second, the decision problem of the PM has a normative nature, whereas the participation decisions by individuals are of positive one. Hence, the socially agreeable objective of the PM does not need to conform with the aggregation of the individual utilities. This is precisely the case of the PM adopting the sufficientarian objective. In this case, the individuals with outcome y greater than y^* may still prefer a higher outcome. However, the PM is not concerned about such preference, once the outcome is above a threshold y^*. Furthermore, as noted by Alcantud, Mariotti, and Veneziani (2022), choosing an appropriate threshold y^* is a task far from obvious, because heterogeneous utilities are hard to compare between people. Third, in our framework, the PM does not have information on the utilities, nor any data to make inference on them. In this setting, focusing on the people with lowest utilities may trivialize the policy learning problem, while considering a weighted average of utilities is operationally tantamount to assuming the knowledge of the distribution of the utilities. Thus we do not find either of these approaches practically attractive as compared to taking the PM's target quantity simply as a known function of the outcome that can be measured.

2.2. Transfer from the Source Population

Our main departure from most existing literature of policy learning is that we consider an *ex ante* policy learning setting where the target population has not yet implemented the policy.

As a consequence, we do not observe the outcomes from the policy in the target population. However, there is a source population which has experimented with the policy and the outcome data from the policy is available. Thus, the policy learning problem is summarized as one in which the PM searches for an optimal policy for the target population using data generated from the source population.

We assume that the source population and the target population are different, and hence not all the information on the experiment setting in the source population is relevant to the PM. Here, we clarify the available data for the PM regarding the source population and the transferable aspects of the experiment setting.

We define the conditional CDF of the potential outcomes and the conditional average total cost (CATC) as follows: for $d = 0, 1$,

$$F_d(y \mid x) = P\{Y(d) \leq y \mid X = x\} \text{ and } c(x) = \mathbf{E}[C_O + C_T \mid X = x], \;\; x \in \mathcal{X}.$$

We assume that $(\tau(\cdot), c(\cdot))$ is identified from the source population and transferable to the target population.

Assumption 2.1 (Identification and Transferability)**.** For each $x \in \mathcal{X}$, $(F_1(\cdot \mid x), F_0(\cdot \mid x), c(x))$ is identified in the source population and identical between the source and target populations.

It is well known that the identification of the conditional CDFs, F_d, can be obtained under the unconfoundedness condition: $(Y(1), Y(0)) \perp\!\!\!\perp D \mid X$. This condition excludes partial compliance in the experiment in the source population. We will later discuss extension of our framework to the case with partial compliance. The assumption also requires a transferability condition that requires $(F_1(\cdot \mid x), F_0(\cdot \mid x), c(x))$ to remain the same as we move from the source to target population. However, we assume that all other aspects of the source population are not transferable to the target population. For example, the distribution of the covariates can be different between the two populations, and hence the average treatment effect (ATE) is not transferable between the populations.

A major challenge for the PM in this setting is that the incentives for participating in the treatment program to differ across the two populations. The difference arises naturally as participation decisions can involve various social, cultural factors that are distinct across different populations. Due to the difference in incentives for treatment participation, we cannot use, for example, the propensity scores estimated from the source population to predict individuals' participation behavior in the target population.[3]

[3]However, the propensity score from the source population can still be useful for identifying $\tau(x)$ which is assumed to be transferable to the target population.

2.3. Stochastic-Dominance Characterization of Policy Learning

Given the budget constraint setting, an exclusive focus on the brute-force worst possible scenario can lead to a highly conservative decision, if the collection of probabilities $\mathcal{P}$ and the class of utility functions $\mathcal{U}$ under consideration is overly large. For example, the maximin solution can be trivial, such as treating no one, if there is any slightest harm expected for a certain group that extremely sensitive to any possible harm from the treatment. To alleviate this issue of conservativeness, we introduce additional, mild conditions for the probabilities and utilities so that the optimal solution is reasonable, admits an explicit form, and is practically implementable using the data from the source population.

2.3.1. **Compliance Guarantee.** First, we place restrictions on $\mathcal{P}$ by introducing the following independence condition.

Assumption 2.2 (Utilities and Costs)**.** For the utilities and costs in the target population, we assume that the following conditions are satisfied.

(i) $(Y(1), Y(0)) \perp\!\!\!\perp V_{-X} \mid X$.

(ii) $(C_O, C_T) \perp\!\!\!\perp V_{-X} \mid X$.

Assumption 2.2(i) says that the potential outcomes are dependent with unobservable component of the payoff only through the covariates X. This seems reasonable when the effectiveness of the treatment outcomes are less related to the idiosyncratic component of individual characteristics and are transferable to any individual with the same covariate group. An extreme example would be the effect of medical treatment which is scientifically verified so that it is deemed transferable to a wide population of the same observable category. In this case, the difference in outcome for the same policy across different populations is likely to be from the different participation incentives. Condition (ii) says that the offer cost and treatment cost are not related to the unobserved idiosyncratic payoff component given the observed covariates. These costs are most likely to arise from the cost of administration and implementation of the treatment from the PM side. (Note that the individual cost in participating the program is subsumed in the payoff U (hence, U is a net payoff), and not included in C_O, C_T.) Hence, this condition seems plausible in practice.

Under Assumption 2.2, we show that the policy learning problem (2.4) can reformulated as one under compliance guarantee. To formalize this result, it is convenient to introduce the following definition: for map $f : \mathcal{X} \to \mathbf{R}$,

$$Q(g; f) = \mathbf{E}[g(X) f(X)].$$

We define the conditional average offer cost and the conditional average treatment cost:

$$c_O(x) = \mathbf{E}[C_O \mid X = x] \text{ and } c_T(x) = \mathbf{E}[C_T \mid X = x].$$

Recall that a_P denotes the individual's participation (or acceptance) rule which depends on the expected payoff from the treatment. Define the propensity score,

$$\pi_u(x) = \mathbf{E}[a(V;u) \mid X = x],$$

and the conditional average treatment effect (CATE) (in terms of $q(Y(d),X)$):

$$\tau(x) = \mathbf{E}[q(Y(1),X) - q(Y(0),X) \mid X = x].$$

Conditional independence in Assumption 2.2 allows us to rewrite (2.3) as:

$$Q_B(g,u;\tau,\mathbf{c}) = \begin{cases} Q(g;\tau\pi_u), & \text{if } Q(g;c_O) + Q(g;c_T\pi_u) \leq B \\ -\infty, & \text{otherwise,} \end{cases} \tag{2.6}$$

where $\tau(x)$ denotes CATE as defined above and $\mathbf{c} = (c_O, c_T)$. (Here, 1 denotes the constant function taking value number one.) The policy learning problem thus becomes:

$$\sup_{g \in \mathcal{G}} \inf_{u \in \mathcal{U}} Q_B(g,u;\tau,\mathbf{c}). \tag{2.7}$$

Our budget constraint in (2.6) suggests that we focus only on the "least favorable" subjects who are at the margin in a certain sense. We say that a policy is ***harmless***, if the distribution of the outcome after treatment stochastically dominates the distribution of the outcome before the treatment. Then the least favorable subjects are those who participates in the program if and only if the program is harmless for them. So, a program that guarantees full compliance is one that is harmless.

2.3.2. **Almost Stochastic Dominance.** The solution to (2.7) can be conservative depending on the class of utility functions that are allowed in the target population. Define $\overline{\mathcal{H}}$ to be the set of all non-decreasing, differentiable functions. For each $\epsilon \in (0,1]$, let[4]

$$\mathcal{H}_\epsilon := \left\{ h \in \overline{\mathcal{H}} : \sup_{t \in [-M,M]} h'(t) \leq \left(\frac{2-\epsilon}{\epsilon} \right) \inf_{t \in [-M,M]} h'(t) \right\}.$$

Then, when $\epsilon = 0$, we take $\mathcal{H}_0$ to be the set of non-decreasing functions on $[-M,M]$. We can think of the class $\mathcal{H}_\epsilon$ as a set in which the random utility $u(\cdot,V)$ realizes. The class $\mathcal{H}_\epsilon$ entails a restriction on how the marginal random utility varies from a low outcome to high outcome.

[4]Without imposing restrictions on the space of utility functions, the maximin approach that we propose later becomes trivial, because the worst scenario is achieved by considering strictly decreasing utility functions for an individual $X = x$ with $\tau(x) > 0$. Then, no one has an incentive to participate and the optimal treatment assignment is to deny treatment to everybody.

The restriction gets stronger when ϵ is larger, so that the classes $\mathcal{H}_\epsilon$, $\epsilon \in [0,1]$, are nested: if $\epsilon_1 < \epsilon_2$, then $\mathcal{H}_{\epsilon_2} \subseteq \mathcal{H}_{\epsilon_1}$.

The connection between $\mathcal{H}_\epsilon$ and stochastic dominance constraints is established by Leshno and Levy (2002) and plays a crucial role in our setting. Let $F_0(\cdot \mid x)$ and $F_1(\cdot \mid x)$ denote the conditional CDF of control and treated outcomes, respectively, given covariate $x \in \mathcal{X}$. Define the degree of AFOSD at x as:

$$\delta(x) = \begin{cases} \dfrac{\int \left[F_1(y \mid x) - F_0(y \mid x)\right]^+ dy}{\int |F_1(y \mid x) - F_0(y \mid x)| dy}, & \text{if denominator is positive,} \\ 0, & \text{otherwise.} \end{cases}$$

According to Leshno and Levy (2002), we have the equivalence:

$$\delta(x) \leq \epsilon \iff \mathbf{E}[h(Y(1)) \mid X = x] \geq \mathbf{E}[h(Y(0)) \mid X = x] \text{ for all } h \in \mathcal{H}_\epsilon. \tag{2.8}$$

For $\epsilon \in [0, 1/2]$, we define

$$\mathcal{X}_\epsilon = \left\{x \in \mathcal{X} : \delta(x) \leq \epsilon\right\}.$$

The set $\mathcal{X}_\epsilon$ consists of the covariate groups such that the degree of AFOSD is bounded by ϵ.

We introduce the following assumption on $\mathcal{U}$ on the utility functions in the target population. Let $\tilde{\mathcal{U}}$ be the class of real valued measurable functions on $\mathcal{Y} \times \mathcal{V}$.

Assumption 2.3. For the target population, there exists $\epsilon \in [0, 1/2]$ such that

$$\mathcal{U} = \{u \in \tilde{\mathcal{U}} : u(\cdot, v) \in \mathcal{H}_\epsilon \text{ for all } v \in \mathcal{V}\}.$$

This assumption says that $\mathcal{U}$ has to be flexible enough, so that each individual in the target population has random utility that can take the value of any element in $\mathcal{H}_\epsilon$ for some ϵ. We assume that the PM knows this constant ϵ. This assumption leads to simplification in the budget constraint in (2.6).

Lemma 2.1. *Suppose that Assumptions 2.3 holds. Then,*

$$\inf_{u \in \mathcal{U}} Q_B(g, u; \tau, \mathbf{c}) = \inf_{u \in \mathcal{U}} Q_B(g, u; \tau, (c_O + c_T, 0)), \quad \mathbf{c} = (c_O, c_T).$$

Note that $Q_B(g, u; \tau, (c_O + c_T, 0))$ involves the acceptance rule a only in the objective function, not in the budget constraint. This is because budget violation outweighs all benefits, and hence the budget constraint a policy has to adhere to is the one under full compliance, which involves $c_O + c_T$. Hence, there is no differentiating between c_O and c_T in this setting. We define

$$c = c_O + c_T.$$

We can interpret $c(x)$ as the group level marginal treatment cost when the group x complies with the treatment. Thus, the PM chooses a policy among those that ensures full compliance.

The following lemma shows that the budget-constrained expected outcome also takes a simple form that involves the potential outcomes and costs only through $\tau(x)$ and $\delta(x)$.

Lemma 2.2. *Suppose that Assumptions 2.1-2.3 hold for some $\epsilon \in [0,1/2]$. Suppose further that a policy g satisfies that $\mathbf{E}[c(X)g(X)] \leq B$. Then,*

$$\inf_{u \in \mathcal{H}_\epsilon} Q_B(g,u;\tau,\mathbf{c}) = \mathbf{E}\Big[g(X)\tau(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\Big] + \mathbf{E}\Big[g(X)\tau(X)\mathbf{1}\{\tau(X) < 0\}\Big].$$

This lemma plays a crucial role for characterizing the optimal treatment assignment. It shows that the adversarial welfare is completely characterized by τ and δ only.

2.4. Optimal Policy

2.4.1. **An Explicit Form.** Our main result is to provide an explicit characterization of the optimal solution g^* in (2.5). We now describe a solution to (2.5). Define

$$\rho(x) := \frac{\tau(x)}{c(x)}.$$

Thus $\rho(x)$ is the CATE relative to the cost. We can view it as the inverse of Cost-Effectiveness-Ratio (CER) used in the health care evaluations.[5] We call $\rho(x)$ the ***Returns-to-Treatment (RTT)***. The result below provides an explicit form of the optimal policy in the target population problem in (2.5).

Theorem 2.1. *Suppose that Assumptions 2.1-2.3 hold. Then the following defines an optimal policy g^* to (2.5):*

$$g^*(x) = \begin{cases} 1 & \text{if } x \in \mathcal{X}_\epsilon \text{ and } \rho(x) > k^*, \\ r & \text{if } x \in \mathcal{X}_\epsilon \text{ and } \rho(x) = k^*, \\ 0 & \text{otherwise,} \end{cases}$$

where the threshold k^* is defined as

$$k^* = \inf\Big\{k \geq 0 : \mathbf{E}\Big[\mathbf{1}\{X \in \mathcal{X}_\epsilon, \rho(X) > k\}c(X)\Big] \leq B\Big\},$$

while r is determined as follows: if $k^*>0$, then r solves

$$\mathbf{E}[\mathbf{1}\{X \in \mathcal{X}_\epsilon, \rho(X) = k^*\}c(X)]r = B - \mathbf{E}\Big[\mathbf{1}\{X \in \mathcal{X}_\epsilon, \rho(X) > k^*\}c(X)\Big],$$

and if $k^* = 0$, then r can be chosen as 0.

[5]There is a debate about the use of ICER (Incremental Cost-Effectiveness-Ratio in health care policies.)

The threshold k^* represents the minimum threshold k such that when we assign to treatment any group with x such that the RTT is above k^*, the total cost does not exceed the budget B. The optimal policy simply suggests that we treat any group x such that the RTT is above this threshold and does not treat the group x if the RTT is below the threshold.

The solution to (2.5) is not necessarily unique. In the characterization of Lemma 2.2, since assigning $g(x) = 1$ for $\tau(x) > 0$ generates adversarial welfare only if $x \in \mathcal{X}_\epsilon$, if the budget constraint is not binding for the solution given by Theorem 2.1, we can set $g(x) = 1$ for some $x \notin \mathcal{X}_\epsilon$ with $\tau(x) > 0$ such that the budget constraint is still met. While this will attain the same adversarial welfare, it may attain higher actual welfare, as we could be overly conservative. We could think about add covariates that are not in $\mathcal{X}_\epsilon$, but has a high treatment to cost ratio.

2.4.2. **Consistent Estimation.** Given the solution form in Theorem 2.1, we consider the sample analogue of the optimal policy using data from the target and source population. For the data requirements on the target population, we observe the random sample of the covariate vectors $X_1, \ldots, X_{n_T}$.

Assumption 2.4 (Random Sample of Covariates in the Target Population)**.** X_i, $i = 1, \ldots, n_T$, are i.i.d. random vectors from the target population.

As for the source population, we assume that we can construct estimators of F_1, F_0, τ and c, denoted by $\hat{F}_1$, $\hat{F}_0$, $\hat{\tau}$ and $\hat{c}$ respectively, based on i.i.d. data $(X_i, Y_i, D_i)_{i=1}^{n_S}$, where n_S denotes the size of the random sample from the source population. For these estimators, we make the following high level assumption.

Assumption 2.5 (Uniform Convergence)**.** (i) There exists $\alpha > 0$ such that as $n_S \to \infty$,

$$\sup_{x \in \mathcal{X}, y \in \mathcal{Y}} \left| \hat{F}_d(y \mid x) - F_d(y \mid x) \right| = O_p(n_S^{-\alpha}), \text{ for } d = 0, 1,$$

$$\sup_{x \in \mathcal{X}} |\hat{\tau}(x) - \tau(x)| = O_p(n_S^{-\alpha}) \text{ and}$$

$$\sup_{x \in \mathcal{X}} |\hat{c}(x) - c(x)| = O_p(n_S^{-\alpha}).$$

(ii) $\underline{c} \leq c(X) \leq \bar{c}$ for some $\underline{c} > 0$ with probability one.

(iii) Conditioned on it being positive, the random variable $\rho(X)$ admits a density that is bounded above on its support.

Assumption 2.5(iv) eliminates the necessity of a tie breaking rule. An optimal policy is then

$$g^*(x) = \mathbf{1}\{x \in \mathcal{X}_\epsilon, \rho(x) > k^*\},$$

where k^* is as defined in Theorem 2.1. We write

$$\delta(x; F_1, F_0) = \delta(x),$$

making the dependence of $\delta(x)$ on F_0 and F_1 explicit. Since all the objects in g are not known, we use a plug-in estimator of g. Define

$$\hat{\mathcal{X}}_{\epsilon_{n_S}} = \left\{x \in \mathcal{X} : \delta\left(x; \hat{F}_1, \hat{F}_0\right) \leq \epsilon_{n_S}\right\},$$

which is an estimator of the set $\mathcal{X}_{\epsilon_{n_S}}$ for some sequence $\epsilon_{n_S} \to \epsilon$.

First, we construct the estimated threshold $\hat{k}$ is given by

$$\hat{k} = \inf\left\{k \geq \xi_{n_S} : \frac{1}{n_T}\sum_{i=1}^{n_T} \mathbf{1}\left\{X_i \in \hat{\mathcal{X}}_{\epsilon_{n_S}}, \hat{\rho}(X_i) > k\right\} \leq B\right\}, \tag{2.9}$$

where ξ_n is bounded sequence which converges to zero at a certain rate that we specify later. Define

$$\hat{g}(x) := \mathbf{1}\left\{x \in \hat{\mathcal{X}}_{\epsilon_{n_S}}, \hat{\rho}(x) > \hat{k}\right\}.$$

The following theorem establishes the consistency of $\hat{k}$.

Theorem 2.2. *Suppose that Assumption 2.5 holds. Suppose further that conditional on $\rho(X) > 0$, the random variable $\rho(X)$ has continuous density that is bounded above zero. Let ξ_{n_S} be a bounded sequence such that $n_S^{\alpha}\xi_{n_S}^2 \to \infty$ for some $\alpha > 0$ in Assumption 2.5 (ii), and let $\epsilon_{n_S} \to \epsilon$ be monotone decreasing satisfying $n_S^{\alpha}\xi_{n_S}^2|\epsilon_{n_S} - \epsilon| \to \infty$. Then,*

$$\hat{k} \to_p k^*,$$

as $n_S, n_T \to \infty$.

The following theorem establishes the consistency of the estimated policy $\hat{g}$.

Theorem 2.3. *Suppose that the conditions of Theorem 2.2 holds. Then,*

$$\int \mathbf{1}\{x : \hat{g}(x) \neq g^*(x)\} dP_X(x) \to_p 0,$$

as $n_S, n_T \to \infty$.

The theorem says that the probability of the covariate groups on which the estimated policy and the true optimal policy differing goes to zero as $n \to \infty$.

2.4.3. **Practical Implementation.** The procedure described in Section 2.4.2 can be viewed as constructing a set C_{n_S}, defined as

$$C_n = \left\{x \in \hat{\mathcal{X}}_{\epsilon_{n_S}} : \hat{\rho}(x) > \xi_{n_S}\right\},$$

with the goal of satisfying the following two properties: as $n_S \to \infty$,

$$P(\left\{x \in \mathcal{X}_\epsilon : \rho(x) > \xi_{n_S}/2\right\} \subseteq C_{n_S}) \to 1, \tag{2.10}$$

and

$$P(X \in \{x \in \mathcal{X}_\epsilon^C : \rho(x) > \xi_{n_S}/2\} \cap C_{n_S}) \to 0. \tag{2.11}$$

In words, C_{n_S} includes valid covariates with large enough RTT, and excludes invalid covariates appropriately. Given these properties, an empirical rule can be obtained by restricting treatment offers to covariates in C_{n_S}, with the RTT cutoff estimated accordingly.

The tuning parameters ξ_{n_S} and ϵ_{n_S} play crucial roles. For each covariate x, when the conditional CDFs are not identical across treatment and control, $\delta(x)$ is defined as

$$\delta(x) = \frac{\int \left[F_1(y \mid x) - F_0(y \mid x)\right]^+ dy}{\int |F_1(y \mid x) - F_0(y \mid x)| dy},$$

and it is possible for both the denominator here and $\delta(x)$ itself to be arbitrarily small. This means $\delta(x, \hat{F}_1, \hat{F}_0)$ may fail to be uniformly consistent across x despite uniform consistency of the conditional CDF estimates. By enforcing that covariates need to satisfy $\hat{\rho}(x) \geq \xi_{n_S}$ to be offered treatment, we impose a positive lower bound on the denominator. If this lower bound goes to 0 at a slow enough rate, $\delta(x, \hat{F}_1, \hat{F}_0)$ will be uniformly consistent across x, and can detect AFOSD violations of size $\epsilon_{n_S} - \epsilon$ given that this does not go to zero too quickly, and leads to (2.10) and (2.11).

Furthermore, if budget B is such that

$$\mathbb{E}[B - \mathbf{1}(X \in \mathcal{X}_\epsilon)\mathbf{1}(\rho(X) > 0)] > 0 \quad \text{but} \quad \mathbb{E}[B - \mathbf{1}(X \in \mathcal{X}_\epsilon)\mathbf{1}(\rho(X) \geq 0)] < 0,$$

and there is a point mass at $\rho(X) = 0$, then

$$\int |\mathbf{1}\{x \in \mathcal{X}_\epsilon : \rho(x) > 0\} - \mathbf{1}\{x \in \mathcal{X}_\epsilon : \hat{\rho}(x) > 0\}| dP_X(x)$$

may not converge to 0 in probability, as we could be offering treatment to covariates with zero conditional average treatment effect, and budget violation may not go to 0 in probability. This can be dealt with by again requiring that $\hat{\rho}(x) > \xi_{n_S}$ for x to be considered for treatment offer: this is the same as only offering treatment to x which displays strong evidence of positive effects since ξ_{n_S} will be of a slow rate than $\hat{\rho}$.

Based on Theorem 2.2, a feasible choice of tuning parameters of the form

$$\epsilon_{n_S} = \epsilon + c_1\left(\frac{\log n_S}{n_S}\right)^{\alpha/2} \quad \text{and} \quad \xi_{n_S} = c_2\left(\frac{\log n_S}{n_S}\right)^{\alpha/4}$$

for some positive constants c_1 and c_2.

Remark 2.1. *If $\rho(X)$ has no point mass in a neighborhood around 0, we should be able to simply set $\xi_{n_S} = 0$ and set a faster rate for $\epsilon_{n_S} \to \epsilon$ such as*

$$\epsilon_{n_S} = \epsilon + c_1 \left(\frac{\log n_S}{n_S} \right)^{\frac{4\alpha}{5}}.$$

While the empirical rule may mistakenly offer treatments to covariates which violate the AFOSD constraint due to the denominator in $\delta(x)$, which is $\|F_1(\cdot|x) - F_0(\cdot|x)\|_{L_1}$, being too close to 0, such x must have $\rho(x)$ close to 0, and overall that proportion goes to 0 in probability.

Similarly, if the budget constraint is known to be binding, we can also set ξ_{n_S} to be any arbitrary non-negative constant smaller than k^*, including 0, and set a faster rate for $\epsilon_{n_S} \to \epsilon$. Even though C_n may contain covariates $x \notin \mathcal{X}_\epsilon$ due to denominator in $\delta(x)$ being close to 0, such x would never be considered for treatment offer anyway due to the budget constraint.

2.5. Confidence Band Approaches

The key goal of constructing C_n is to ensure that the conditions (2.10) and (2.11) are satisfied. An alternative approach instead defines

$$C_n = \left\{ x \in \mathcal{X} : L(x) \le \epsilon, \hat{\rho}(x) > \xi_{n_S} \right\}.$$

where $L(x)$ is a lower bound of a uniform confidence band for $\delta(x)$. Two common methods of constructing $L(x)$ are:

(1) Bootstrap supremum: Draw bootstrap samples for $\sup_{x:\hat{\rho}(x)>\xi_{n_S}} \left| \hat{\delta}(x) - \delta(x) \right|$ to estimate its q_n quantile (where $q_n \to 1$), denoted $\hat{q}_n$. Then define

$$L(x) = \hat{\delta}(x) - \hat{q}_n.$$

(2) Studentized bootstrap supremum: Estimate pointwise standard errors $\hat{\sigma}(x)$ via bootstrap, then estimate the q_n quantile (where $q_n \to 1$) of

$$\sup_{x:\hat{\rho}(x)>\xi_{n_S}} \left| \frac{\hat{\delta}(x) - \delta(x)}{\hat{\sigma}(x)} \right|,$$

denoted $\hat{q}_n$, and define

$$L(x) = \hat{\delta}(x) - \hat{q}_n \hat{\sigma}(x)$$

To ensure the validity of either method, it is likely necessary to restrict the supremum to the set $\{x : \hat{\rho}(x) > \xi_{n_S}\}$. Unless one assumes that the denominator of $\hat{\delta}(x)$,

$$\|F_1(\cdot \mid x) - F_0(\cdot \mid x)\|_{L_1},$$

is uniformly bounded away from zero for all $x \in \mathcal{X}$, uniform consistency of $\hat{\delta}(x)$ may fail. As a result, even if the budget constraint is binding and $\rho(X)$ has no point mass near zero, taking bootstrap suprema over the entire covariate space can behave poorly. The restriction to covariates with sufficiently large estimated RTT, enforced through $\hat{\rho}(x) > \xi_{n_S}$, helps control this issue.

2.6. Numerical Illustration

We illustrate how the optimal policy works, by comparing its performance with that from a naive policy. The naive policy offers treatment to those with the highest Returns-to-Treatment (RTT). It mirrors the approach of Sun et al. (2024), who assume full uptake once treatment is offered. On the other hand, the optimal policy considers the individual's take-up decision, based on the AFOSD (called AFOSD-based policy here). This rule first checks whether the treatment outcome from each potential recipient satisfies the approximate first-order stochastic dominance (AFOSD) criterion up to a tolerance level ϵ_{target}. Then, eligible agents are then ranked by RTT. This policy takes into account the possibility that some individuals may reject the offer of treatment.

We consider different variance configurations for the conditional distribution of the potential outcomes given the covariates. It is not easy to induce highly risk-averse individuals to participate in the treatment, when the treatment is highly effective on average, yet can carry some negative net utilities with a positive probability. Thus, risk-averse people have different incentives to participate in the treatment, if the treatment outcomes differ in their variances, yielding different compliance behaviors.

2.6.1. **Data Generating Process.** First, we generate covariates and control outcomes as

$$X_i \sim \text{Uniform}(0,1) \quad \text{and} \quad Y_i(0) \sim \text{Uniform}(1.2, 1.4)$$

which are independent. Conditional on $X_i = x$, treated outcome $Y_i(1)$ has the mean: $m_1(x) = 1.8 + 0.05x$, but for their variances, we consider two specifications:

$$\textbf{Model A: } \sigma^2_{1,A}(x) = \left[0.7\cos^2\left(\frac{\pi x}{2}\right)\right]^2 \text{ and } \textbf{Model B: } \sigma^2_{1,B}(x) = \left[0.7\cos^2(x-1)\right]^2.$$

In Model A, the CATE and the conditional variance of $Y_i(1)$ align in the sense that the conditional variance declines with x while CATE increases in x. In Model B, the CATE and the conditional variance of stochastic dominance moves in the opposite direction: In both cases, conditional on $X_i = x$, the treated outcomes in both models are normally distributed, but truncated $\pm 2\sigma_1(x)$ around their respective means.

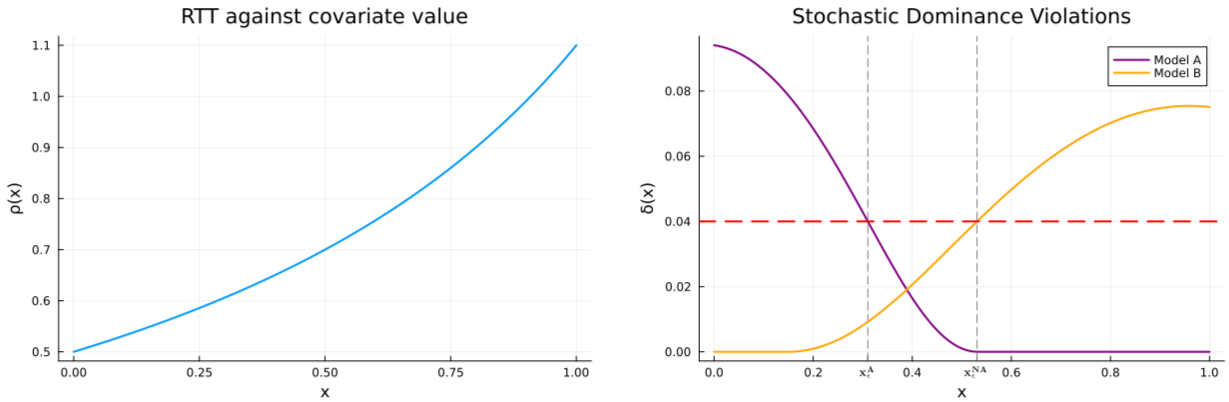


FIGURE 1. Returns-to-Treatment and Stochastic Dominance Violation

Note: The left figure shows the cost Returns-to-Treatment (RTT) against the values of covariates, whereas the right figure plots AFOSD violation against covariates. In our design, the RTT increases with x in both models A and B. However, due to the different behavior of the conditional variance of $Y(1)$ given $X = x$ between models, the violation of the approximate first order stochastic dominance increases with x in Model B whereas it decreases with x in Model A.

Next, conditional on X_i, treatment costs C_i are drawn from Uniform$(0.5-0.5x, 1.5-0.5x)$, so that $\mathbf{E}[C_i \mid X_i = x] = 1-0.5x$.

Agents evaluate outcomes using expected utility from a CRRA utility function

$$h(y;\gamma_i) = \frac{(y-0.45)^{1-\gamma_i}-1}{1-\gamma_i}$$

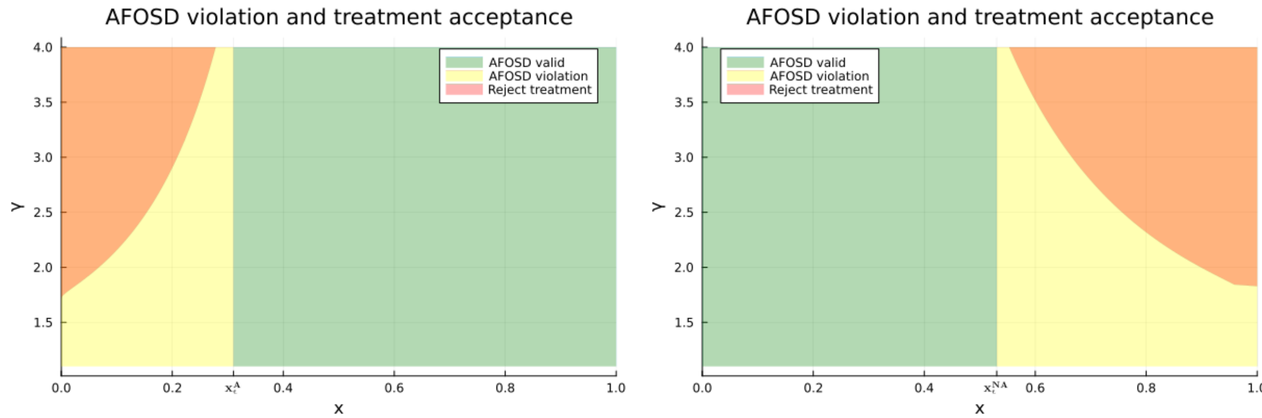
where each $\gamma_i \sim$ Uniform$(1.1, 4)$ is drawn independently of all other variables in both models. Consequently, an agent with (x,γ) will accept a treatment offer if and only if

$$\mathbf{E}[h(Y_i(1),\gamma_i)|X_i = x, \gamma_i = \gamma] \geq \mathbf{E}[h(Y_i(0),\gamma_i)|X_i = x, \gamma_i = \gamma].$$

Both models A and B share the same conditional means for treatment effect and cost. As x increases, the conditional average treatment effect rises, while expected cost falls, so the RTT is increasing in x (see Figure 1(A)).

However, the models differ in how the conditional variance behaves, resulting in opposite patterns of stochastic dominance violation. This leads to a threshold x_ϵ^A such that $\delta(x) \leq \epsilon_{\text{target}}$ if and only if $x \geq x_\epsilon^A$. By contrast, CATE and stochastic dominance do not align in model B, as conditional variance of $Y_i(1)$ increases with x, and $\delta(x) \leq \epsilon_{\text{target}}$ if and only if $x \leq x_\epsilon^B$ for some threshold x_ϵ^B (see Figure 1(b)).

We set the tolerance level at $\epsilon_{\text{target}} = 0.04$. Based on the characterization of δ in Section 2, $\mathcal{H}_{\epsilon_{\text{target}}}$ contains every $h(y;\gamma)$ for γ in the support of γ_i. As a result, any individual with x such that $\delta(x) \leq \epsilon_{\text{target}}$ will accept treatment. On the other hand, when $\delta(x) > \epsilon_{\text{target}}$, whether an individual accepts treatment depends on his utility function. Since higher γ implies



(A) Covariate-utility decision under model A

(B) Covariate-utility decision under model B

FIGURE 2. AFOSD Violation and Treatment Acceptance

Note: In each plot, the green region indicates covariates x for which $\delta(x) \le \epsilon_{\text{target}}$. Agents in the yellow region have $\delta(x) > \epsilon_{\text{target}}$ but are not risk averse enough to reject treatment. This is in contrast to agents in the red region who have $\delta(x) > \epsilon_{\text{target}}$ and are highly risk averse, so that they would reject treatment if offered.

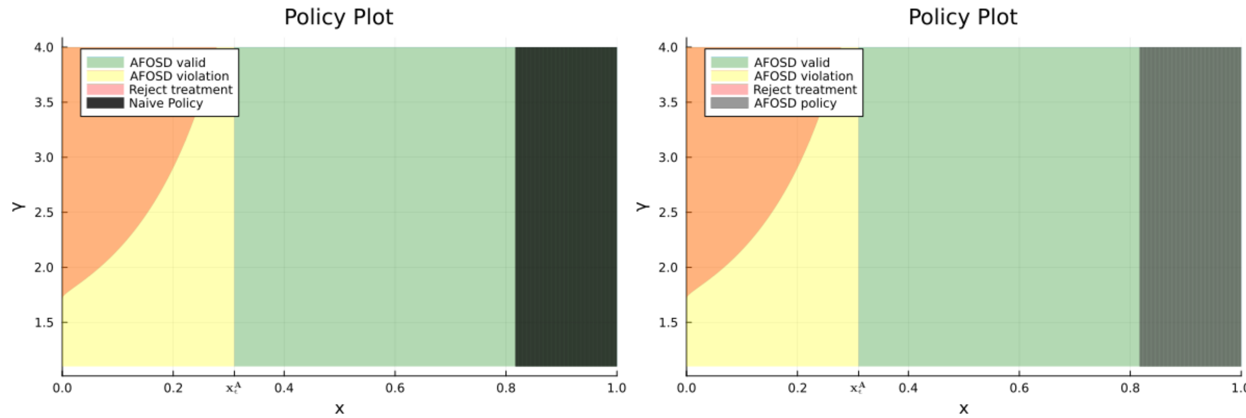
a greater risk aversion and because the conditional variance of $Y(1)$ grows in x, a portion of the population, characterized by high x and γ, will in fact reject treatment. The rest, despite having $\delta(x) > \epsilon_{\text{target}}$, still accepts treatment if offered. Figure 2 illustrates these three regions in the (x, γ) plane for each model, where:

- Green region contains those with $\delta(x) \le \epsilon_{\text{target}}$.
- Yellow region contains those with $\delta(x) > \epsilon$ but would still accept treatment.
- The red region contains those with $\delta(x) > \epsilon_{\text{target}}$ and would actually reject treatment.

2.6.2. **Comparing Population Solutions.** We fix budget at $B = 0.1$, a level too small to offer treatment to everyone even though $\rho(x) > 0$ for all x. Under the naive rule, we simply set a cutoff k_{naive} so that offering treatments to all agents with $\rho(x) > k_{\text{naive}}$ exactly exhausts the budget. By contrast, our AFOSD rule requires two simultaneous conditions for an offer at covariate value x: i) a sufficiently high $\rho(x) > k_{\text{AFOSD}}$, and ii) a small enough AFOSD violation, i.e. $\delta(x) \le \epsilon_{\text{target}}$. The extent to which these two rules agree depend on the joint distribution of potential outcomes.

Model A

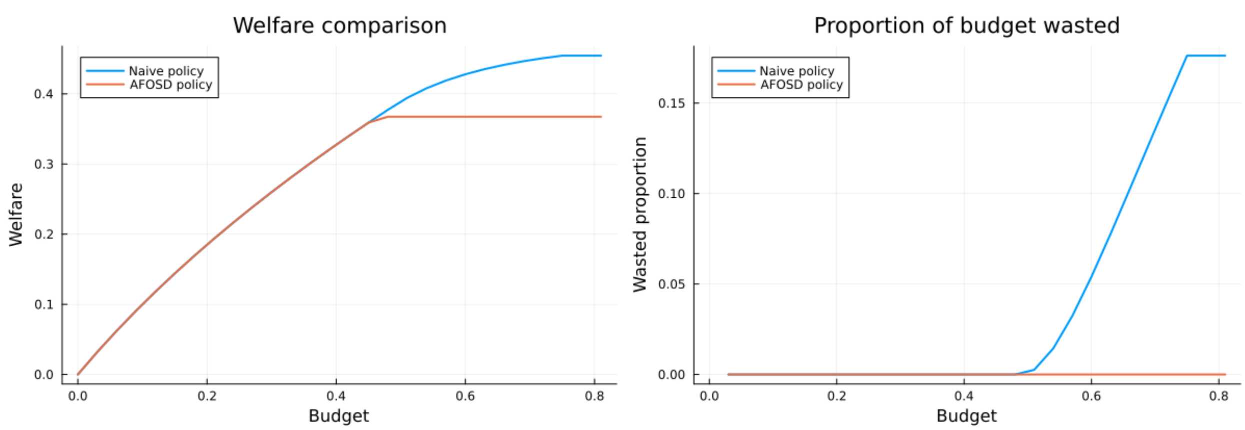
In Model A, high RTT is associated with low AFOSD violation, so for low budgets, both rules target the same subpopulation with high x. As shown in Figure 3, at budget equals 0.1, the naive and AFOSD rules align perfectly, offering treatment to exactly the same set. Because treatment offers are based solely on x, both rules result in vertical slices in the (x, γ) plane.



(A) Agents offered treatment under the naive rule (black region)

(B) Agents offered treatment under the AFOSD rule (grey region)

FIGURE 3. Naive and AFOSD policy under model A

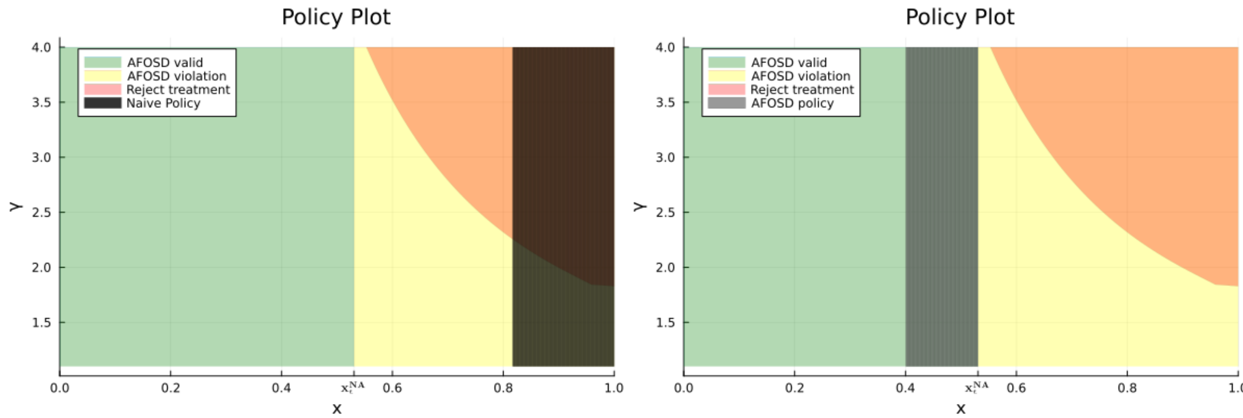


(A) Welfare under naive and AFOSD treatment rules as budget varies.

(B) Proportion of budget wasted on rejected offers under naive and AFOSD treatment rule as budget varies.

FIGURE 4. Welfare and wastage against budget under model A

As the budget increases, the policies diverge. The naive rule ignores stochastic dominance violations and continues expanding offers until all x are included, regardless of compliance. The AFOSD rule, however, stops expanding once all covariates satisfying $\delta(x) \leq \epsilon_{\text{target}}$ (i.e. $x \geq x_{\epsilon}^{A}$) are included; it excludes the remaining x values that violate the AFOSD constraint. This difference is visible in Figure 4, which shows welfare and budget wastage as functions of budget. At higher budgets, AFOSD treats a strict subset of those offered by the naive rule, resulting in lower total welfare but zero wastage, since every dollar goes to an agent who accepts.



(A) Agents offered treatment under the naive rule (black region)

(B) Agents offered treatment under the AFOSD rule (grey region)

FIGURE 5. Naive and AFOSD policy under model NA

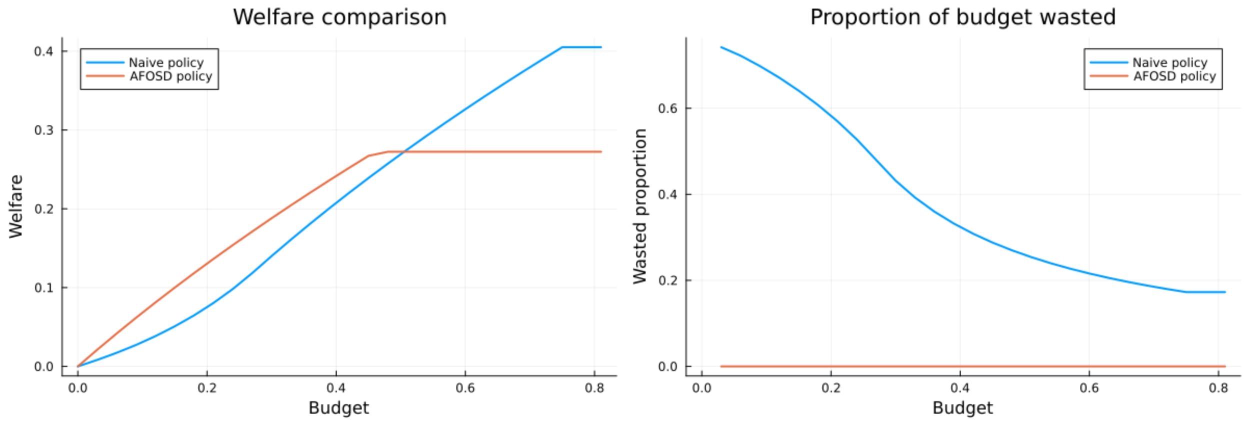
Model B

Model B presents a stark contrast. Here, higher RTT is correlated with greater stochastic dominance violation. The naive policy starts at the maximum $x = 1$ and expand leftwards until the budget is depleted as before, whereas the AFOSD decision rule starts expanding from $x = x_\epsilon^B$ instead of $x = 1$, thereby reserving offers for those covariate values where risk-averse agents are guaranteed to accept.

Figure 5 highlights a key weakness of the naive policy: by targeting the highest RTT values without regard to compliance in model B, it ends up offering treatment to many agents who will refuse. For example, at $x = 1.0$, roughly 75 percent of individuals possess a γ large enough to reject treatment. Although the naive rule maximizes allocation based on RTT, this translates poorly into actual welfare gains once refusal is accounted for.

In contrast, the AFOSD rule confines offer to the region where $\delta(x) \leq \epsilon_{\text{target}}$. While this constraint may exclude some high RTT individuals, every treatment offered under AFOSD rule is taken up, ensuring that each allocation fully contributes to welfare. Consequently, AFOSD rule yields more reliable welfare improvements by aligning allocation decisions with agents' true willingness to accept treatment.

Figure 6(a) demonstrates that, in model B, targeting only individuals with a compliance guarantee is advantageous at low to medium budget levels. As the budget increases, the number of agents offered treatment under the AFOSD rule rises, leveling off at a budget of approximately 0.44, once all agents who satisfy the AFOSD criterion are included. Up to this threshold, the compliance guarantee property of AFOSD results in higher welfare than the naive rule. Beyond this point, however, only the naive rule continues to expand the offer set, until reaching budget level of around 0.75, at which stage everyone is offered treatment and welfare



(A) Welfare under naive and AFOSD treatment rules as budget varies

(B) Proportion of budget wasted due to rejection under naive and AFOSD treatment rule as budget varies

FIGURE 6. Welfare and wastage against budget under model B

plateaus. Due to this difference in targeting, the naive rule ultimately overtakes AFOSD at high budget levels, capturing additional high RTT agents who may not always accept treatment.

This pattern is mirrored in Figure 6(b), which breaks down wastage. Under the naive policy, a nontrivial share of spending goes to agents who reject treatment, and is especially high for low budget levels, whereas AFOSD incurs zero wasted expenditure.

2.6.3. **With Estimation.** To assess finite-sample performance, we estimate each policy in Figure 3 and 5 corresponding to $B = 0.1$, and examine how their welfare gap evolves as sample size increases. The estimation procedure is as follows:

(1) **Naive policy estimation**

Let

$$\hat{\rho}(x) = \frac{\hat{\tau}(x)}{\hat{c}(x)}$$

where $\hat{\tau}, \hat{c}$ are estimates for CATE and conditional mean cost using the Nadaraya Watson estimator. The naive cutoff is then obtained by

$$\hat{k}_{naive} = \sup_k \left\{ \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\hat{\rho}(X_i) \geq k)\hat{c}(X_i) - B \geq 0 \right\}$$

which leads to the estimated naive policy

$$\hat{g}_{naive}(x) = \mathbf{1}\big(\hat{\rho}(x) > \hat{k}_{naive}\big)$$

(2) **AFOSD policy estimation**

$\hat{\rho}(x)$ is defined as above. We also estimate the AFOSD violation using

$$\hat{\delta}(x) = \frac{\frac{1}{M}\sum_{j=1}^{M}\left[\hat{F}_1(y_j|x) - \hat{F}_0(y_j|x)\right]^+}{\frac{1}{M}\sum_{j=1}^{M}\left|\hat{F}_1(y_j|x) - \hat{F}_0(y_j|x)\right|}$$

where $\{y_1, \cdots, y_M\}$ is a fine grid of y values in the support of Y. With this, we obtain lower bound $L(x)$ of a uniform confidence band for $\delta(x)$ using studentized bootstrap as outlined in Section 2.5, and solve for the threshold as

$$\hat{k}_{\text{AFOSD}} = \sup_k \left\{ \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\hat{\rho}(X_i) \geq k)\mathbf{1}\big(L(X_i) \leq \epsilon_{\text{target}}\big)\hat{c}(X_i) - B \geq 0 \right\}$$

and the policy estimate is

$$\hat{g}_{\text{AFOSD}}(x) = \mathbf{1}\big(\hat{\rho}(x) \geq \hat{k}_{\text{AFOSD}}\big)\mathbf{1}\big(\hat{L}(x) \leq \epsilon_{\text{target}}\big).$$

For each sample size $n \in \{1000, 2000, 3000, 4000\}$, we proceed as follows:

(a) Draw a dataset of size n and then estimate both the naive and AFOSD rules. This is repeated 200 times. In each of these 200 replications, the AFOSD lower-bound threshold is constructed using the studentized bootstrap (method 3) with $q_n = 1 - 0.5n^{-\frac{1}{4}}$.

(b) For each replication $r = 1, \cdots, 200$, and each covariate value x in our grid, we record whether the naive policy $\hat{g}^{(r)}_{\text{naive}}(x)$ and the AFOSD policy $\hat{g}^{(r)}_{\text{AFOSD}}$ offers treatment.

(c) For each point on a fine grid of x values, compute the relative frequency with which each rule would offer treatment:

$$\bar{g}_{\text{naive}}(x) = \frac{1}{200}\sum_{r=1}^{200} \mathbf{1}\left\{\hat{g}^{(r)}_{\text{naive}}(x) = 1\right\}$$

$$\bar{g}_{\text{AFOSD}}(x) = \frac{1}{200}\sum_{r=1}^{200} \mathbf{1}\left\{\hat{g}^{(r)}_{\text{AFOSD}}(x) = 1\right\}$$

These curves represent the "average learned" policies as functions of x.

We then overlay $\bar{g}_{\text{naive}}$ and $\bar{g}_{\text{AFOSD}}$ on the true population policy boundaries to assess their convergence.

2.6.4. **Policy Estimation Under Model A.** Figure 7 shows the average learned policy for each rule in Model A. At budget equal 0.1, the naive and AFOSD policies coincide in both theory and practice. At this budget level, since $\delta(x) \leq \epsilon$ for all x in the shaded area, and the lower confidence band lies below δ with high probability, the estimated policies depend essentially on only $\hat{\rho}$ which tracks ρ reasonably well.

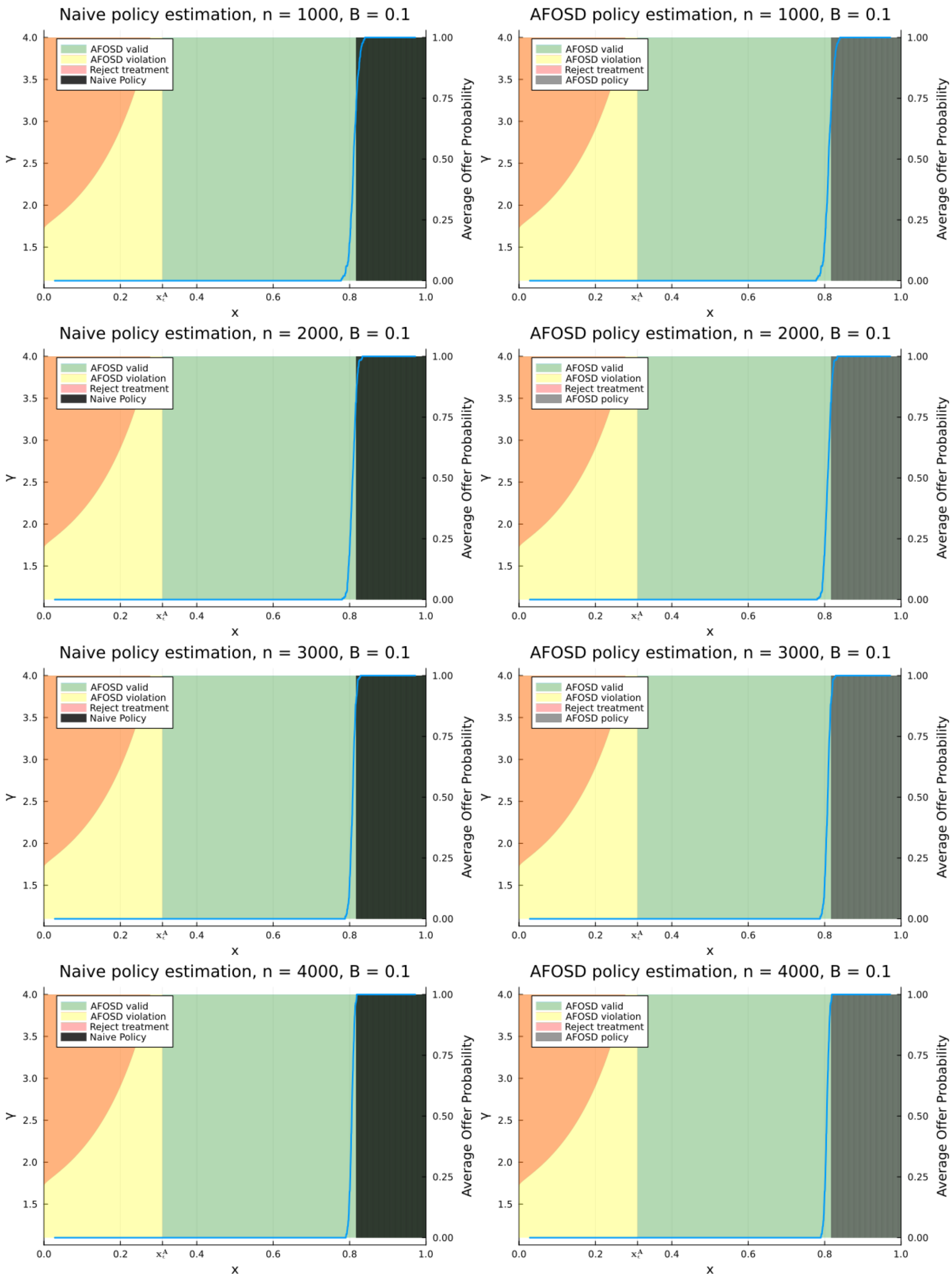


FIGURE 7. Comparison of population and estimated policies under model A

2.6.5. **Policy Estimation Under Model B.** In the left panels of Figure 8, which correspond to the naive policy estimator, the learned policy closely tracks the true population rule even at $n = 1000$. This consistency reflects the fact that ranking by the estimated $\hat{\rho}(x)$ mirrors the ranking by the true $\rho(x)$, so the cutoff procedure reliably selects the same covariate region regardless of sample size.

By contrast, for the estimated AFOSD policy to work well in model B, it has to be able to eliminate points that violate the stochastic dominance constraints, and this is challenging. The right panels in Figure 8, which correspond to the AFOSD policy estimates, reveal noticeable discrepancies for $n = 1000$ and 2000. Although the RTT ranking remains accurate, the one-sided confidence lower bound $\hat{L}(x)$ significantly understates the true AFOSD violation $\delta(x)$, causing the estimator to offer treatment at $x > x_\epsilon$ when it should not. As n grows, however, $\hat{L}(x)$ converges more tightly to $\delta(x)$, and the estimated policy increasingly refrains from offering treatment beyond the true threshold. This shift is evident in the declining "overshoot" of the average offer probability and the gradual alignment of its peak with x_ϵ.

3. Extensions

3.1. Partial Compliance in the Source Population

We have assumed that the source population complied with the program fully, so that the CATE $\tau(x)$ is identified. We can extend our framework to the case where the source population complies only partially. In this case, we can partially identify CATE $\tau(x)$ as an interval, where the interval consists of local average treatment effects, and use the lower bound instead of $\tau(x)$ in our framework.

3.2. Soft Budget Constraints

We have assumed hard budget constraint, where no violation of the budget constraint is allowed. However, in reality, there might be some flexibility in budget violation. We can modify the procedure depending on the degree of violation that the PM is willing to tolerate.

4. Concluding Remarks

We consider a policy learning problem under a budget constraint, where the setting of source population is only partially transferable to the target population. More specifically, the conditional average treatment effects and the conditional cost functions are transferable, whereas the participation incentives are not. By using the almost first order stochastic dominance

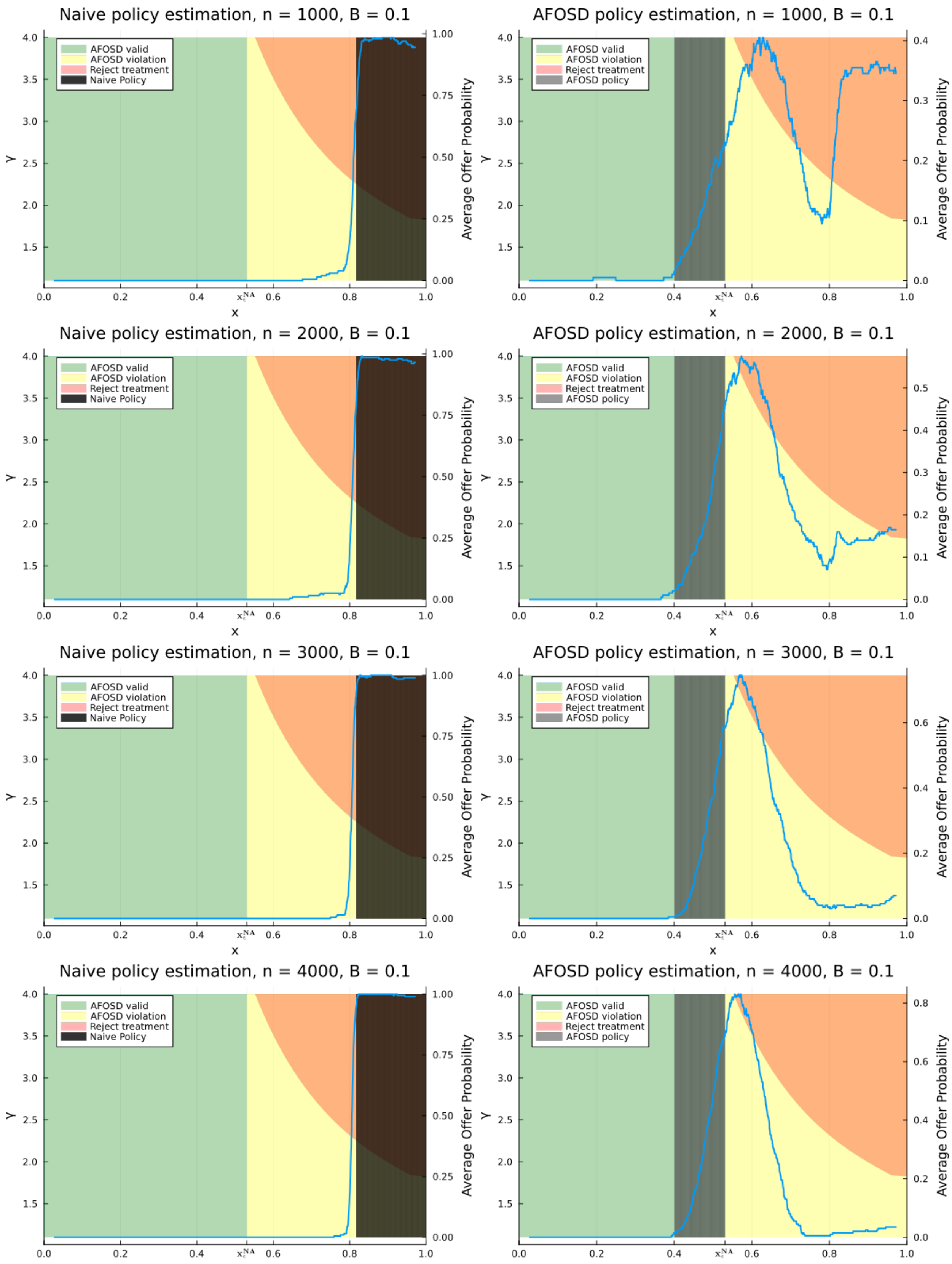


FIGURE 8. Comparison of population and estimated policies under model B

of Leshno and Levy (2002), we find an explicit optimal policy, which assigns treatment only to those with full compliance, with priorities given to those with highest returns to the treatment. We provide an estimator of the optimal policy and establish its consistency.

5. Appendix: Mathematical Proofs

Proof of Lemma 2.1: If $Q(g; c_O + c_T) > B$, then the choice of u which satisfies $u(y, V) = 1$ for all y, indeed being a member of $\mathcal{H}_\epsilon$, leads to $\pi_u(X) = 1$ with probability 1 because in this case, the subject is indifferent to the treatment outcomes. Therefore,

$$Q(g; c_O) + Q(g; c_T \pi_u) = Q(g; c_O + c_T) > B.$$

Budget violation means that welfare under this u is

$$Q_B(g, u; \tau, \mathbf{c}) = -\infty = Q_B(g, u; \tau, (c_O + c_T, 0)).$$

On the other hand, if g satisfies $Q(g; c_O + c_T) \le B$, then for any u, we have

$$Q(g; c_O) + Q(g; c_T \pi_u) \le Q(g; (c_O + c_T, 0), 1) \le B,$$

so that

$$Q_B(g, u; \tau, \mathbf{c}) = Q(g; \tau \pi_u) = Q_B(g, u; \tau, (c_O + c_T, 0)).$$

The last equality follows because $Q_B(g, u; \tau, \mathbf{c})$ does not depend on $\mathbf{c}$ as long as $\mathbf{c}$ satisfies the budget constraint. In both cases, we have

$$\inf_{u \in \mathcal{U}} Q_B(g, u; \tau, \mathbf{c}) = \inf_{u \in \mathcal{U}} Q_B(g, u; \tau, (c_O + c_T, 0)).$$

■

Proof of Lemma 2.2: First, note that for each $x \in \mathcal{X}_\epsilon$ and $v = (x, v_{-x})$ with some $v_{-x} \in \mathcal{V}_{-X}$,

$$\begin{aligned}
&\mathbf{E}[u(Y(1), V) \mid V = v] - \mathbf{E}[u(Y(0), V) \mid V = v] \\
&= \int u(y, v) dP_{Y(1)|V}(y \mid v) - \int u(y, v) dP_{Y(0)|V}(y \mid v) \\
&\ge \inf_{v' \in \mathcal{V}} \left\{ \int u(y, v') dP_{Y(1)|V}(y \mid v) - \int u(y, v') dP_{Y(0)|V}(y \mid v) \right\},
\end{aligned}$$

with $P_{Y(d)|V}$ denotes the regular conditional distribution of $Y(d)$ given V. By Assumptions 2.2(ii) and 2.3(i), the last infimum is bounded from below by

$$\begin{aligned}
&\inf_{h \in \mathcal{H}_\epsilon} \{\mathbf{E}[h(Y(1)) \mid V = v] - \mathbf{E}[h(Y(0)) \mid V = v]\} \\
&= \inf_{h \in \mathcal{H}_\epsilon} \{\mathbf{E}[h(Y(1)) \mid X = x] - \mathbf{E}[h(Y(0)) \mid X = x]\} \ge 0,
\end{aligned}$$

where the last inequality follows by (2.8). Therefore,

$$\pi_u(x) = 1 \text{ if } x \in \mathcal{X}_\epsilon. \tag{5.1}$$

Suppose that $Q(g,c) \le B$. Observe that for any $\epsilon \in [0,1/2]$, since the set of linear functions is in $\mathcal{H}_\epsilon$, if x is in $\mathcal{X}_\epsilon$, then it has to be that $\tau(x) \ge 0$. Conversely, if $\tau(x) < 0$, then $x \notin \mathcal{X}_\epsilon$. Thus we can decompose welfare as follows:

$$\begin{aligned}\inf_{u\in\mathcal{U}} Q_B(g;\tau\pi_u) &= \inf_{u\in\mathcal{U}} \mathbf{E}[g(X)\tau(X)\pi_u(X)] \\ &= \inf_{u\in\mathcal{U}} \mathbf{E}[g(X)\tau(X)\pi_u(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\} + g(X)\tau(X)\pi_u(X)\mathbf{1}\{\tau(X)<0\} \\ &\quad + g(X)\tau(X)\pi_u(X)\mathbf{1}\{X \notin \mathcal{X}_\epsilon, \tau(X) > 0\}].\end{aligned} \tag{5.2}$$

Now, by (5.1), the last infimum is written as

$$\begin{aligned}&\mathbf{E}\Big[g(X)\tau(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\Big] \\ &\quad + \inf_{u\in\mathcal{U}} \mathbf{E}\Big[g(X)\tau(X)\pi_u(X)\mathbf{1}\{\tau(X)<0\} + g(X)\tau(X)\pi_u(X)\mathbf{1}\{X \notin \mathcal{X}_\epsilon, \tau(X) > 0\}\Big].\end{aligned} \tag{5.3}$$

Let us focus on the infimum in (5.3). We choose $\tilde{u} \in \mathcal{U}$ as follows. Suppose that $x \notin \mathcal{X}_\epsilon$ and $\tau(x) > 0$. By the contrapositive of (2.8), for any $x \notin \mathcal{X}_\epsilon$, there exists some $h_x \in \mathcal{H}_\epsilon$ such that

$$\mathbf{E}[h_x(Y(1)) \mid X = x] < \mathbf{E}[h_x(Y(0)) \mid X = x]. \tag{5.4}$$

Using this, we define

$$\tilde{u}(\cdot, v) = \begin{cases} h_x(\cdot), & \text{if } v = (x, v_{-x}) \text{ with } x \notin \mathcal{X}_\epsilon \text{ and } \tau(x) > 0 \\ 1, & \text{otherwise.} \end{cases}$$

This is indeed an element of $\mathcal{U}$ by Assumption 2.3(ii), since the mapping $v \to \tilde{u}(\cdot, v)$ is constant in v_{-x}. Under $\tilde{u}$, for $v = (x, v_{-x})$ such that $x \notin \mathcal{X}_\epsilon$ and $\tau(x) > 0$, (5.4) implies that

$$\begin{aligned}&\mathbf{E}[\tilde{u}(Y(1), V) \mid V = v] - \mathbf{E}[\tilde{u}(Y(0), V) \mid V = v] \\ &= \mathbf{E}[h_x(Y(1)) \mid V = v] - \mathbf{E}[h_x(Y(0)) \mid V = v] \\ &= \mathbf{E}[h_x(Y(1)) \mid X = x] - \mathbf{E}[h_x(Y(0)) \mid X = x] < 0,\end{aligned}$$

so that $\pi_{\tilde{u}}(x) = 0$. On the other hand, for $v = (x, v_{-x})$ with $\tau(x) < 0$, we have

$$\mathbf{E}[\tilde{u}(Y(1), V) \mid V = v] = 1 = \mathbf{E}[\tilde{u}(Y(0), V) \mid V = v],$$

and this implies $\pi_{\tilde{u}}(x) = 1$. Consequently, for any $u \in \mathcal{U}$, we have

$$\begin{aligned}&\mathbf{E}\Big[g(X)\tau(X)\pi_u(X)\mathbf{1}\{\tau(X)<0\} + g(X)\tau(X)\pi_u(X)\mathbf{1}\{X \notin \mathcal{X}_\epsilon, \tau(X) > 0\}\Big] \\ &\ge \mathbf{E}\Big[g(X)\tau(X)\mathbf{1}\{\tau(X)<0\}\Big]\end{aligned}$$

$$= \mathbf{E}\Big[g(X)\tau(X)\pi_{\tilde{u}}(X)\mathbf{1}\{\tau(X) < 0\} + g(X)\tau(X)\pi_{\tilde{u}}(X)\mathbf{1}\{X \notin \mathcal{X}_\epsilon, \tau(X) > 0\}\Big]$$

where the inequality follows because g and the indicator functions are non-negative while τ is negative in the first term, and positive in the second. Therefore, the infimum in (5.3) is equal to

$$\mathbf{E}\Big[g(X)\tau(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\Big] + \mathbf{E}\Big[g(X)\tau(X)\mathbf{1}\{\tau(X) < 0\}\Big].$$

This completes the proof. ∎

Proof of Theorem 2.1: We first consider the case where $\mathbf{E}[\mathbf{1}\{X \in \mathcal{X}_\epsilon, \tau(x) > 0\}c(x)] \leq B$. In this case, setting $k^* = 0$ and $r = 0$ is equivalent to the policy

$$g(x) = 1 \iff x \in \mathcal{X}_\epsilon, \tau(x) > 0.$$

This maximizes the first term in (5.3), while ensuring the second term, which is non-positive, is equal to 0.

For the other case where $\mathbf{E}[\mathbf{1}\{X \in \mathcal{X}_\epsilon$ and $\tau(X) > 0\}c(X)] > B$, we have $k^* > 0$. Denote

$$\mathcal{X}_\epsilon^a = \{x : \rho(x) > k^*, x \in \mathcal{X}_\epsilon\}$$
$$\mathcal{X}_\epsilon^b = \{x : \rho(x) = k^*, x \in \mathcal{X}_\epsilon\}$$
$$\mathcal{X}_\epsilon^c = \{x : \rho(x) < k^*, x \in \mathcal{X}_\epsilon\},$$

so that

$$\mathbf{1}\{x \in \mathcal{X}_\epsilon^a\}\rho(x) > k^*\mathbf{1}\{x \in \mathcal{X}_\epsilon^a\}$$
$$\mathbf{1}\{x \in \mathcal{X}_\epsilon^b\}\rho(x) = k^*\mathbf{1}\{x \in \mathcal{X}_\epsilon^b\}$$
$$\mathbf{1}\{x \in \mathcal{X}_\epsilon^c\}\rho(x) < k^*\mathbf{1}\{x \in \mathcal{X}_\epsilon^c\}. \tag{5.5}$$

Let $\bar{g}$ be an optimal policy to (2.5), and

$$\tilde{g}(x) = \begin{cases} \bar{g}(x), & \text{if } x \in \mathcal{X}_\epsilon, \\ 0, & \text{otherwise.} \end{cases}$$

Observe that

(1) since $\bar{g}$ satisfies the budget constraint, $\tilde{g}$ must also satisfy the budget constraint since it does not treat anymore than $\bar{g}$; and
(2) if $\tau(x) < 0$, then $x \notin \mathcal{X}_\epsilon$, so that $\tilde{g}(x) = 0$;

So that by (5.3), $\tilde{g}$ attains the same adverse welfare as $\bar{g}$. Now, define

$$\tilde{\mathcal{X}}_\epsilon^a = \{x : x \in \mathcal{X}_\epsilon, \tilde{g}(x) = 1\}$$
$$\tilde{\mathcal{X}}_\epsilon^b = \{x : x \in \mathcal{X}_\epsilon, \tilde{g}(x) \in (0,1)\}$$

$$\tilde{\mathcal{X}}_\epsilon^c = \{x : x \in \mathcal{X}_\epsilon, \tilde{g}(x) = 0\}$$

We compare the difference in resulting mean outcomes

$$\mathbf{E}[\tilde{g}(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\tau(X)] - \mathbf{E}[g^*(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\tau(X)]$$

by the decomposition

$$\begin{aligned}
&\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^a \cap \mathcal{X}_\epsilon^b\big\}(1-r)\tau(X)\big] + \mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^a \cap \mathcal{X}_\epsilon^c\big\}\tau(X)\big] \\
&\quad - \mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^b \cap \mathcal{X}_\epsilon^a\big\}(1-\tilde{g}(X))\tau(X)\big] + \mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^b \cap \mathcal{X}_\epsilon^b\big\}(\tilde{g}(X)-r)\tau(X)\big] \\
&\quad + \mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^b \cap \mathcal{X}_\epsilon^c\big\}\tilde{g}(X)\tau(X)\big] - \mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^c \cap \mathcal{X}_\epsilon^a\big\}\tau(X)\big] \\
&\quad - r\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^c \cap \mathcal{X}_\epsilon^b\big\}\tau(X)\big] \\
&\le k^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^a \cap \mathcal{X}_\epsilon^b\big\}(1-r)c(X)\big] + k^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^a \cap \mathcal{X}_\epsilon^c\big\}c(X)\big] \\
&\quad - k^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^b \cap \mathcal{X}_\epsilon^a\big\}(1-\tilde{g}(X))c(X)\big] + k^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^b \cap \mathcal{X}_\epsilon^b\big\}(\tilde{g}(X)-r)c(X)\big] \\
&\quad + k^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^b \cap \mathcal{X}_\epsilon^c\big\}\tilde{g}(X)c(X)\big] - k^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^c \cap \mathcal{X}_\epsilon^a\big\}c(X)\big] \\
&\quad - rk^*\mathbf{E}\big[\mathbf{1}\big\{X \in \tilde{\mathcal{X}}_\epsilon^c \cap \mathcal{X}_\epsilon^b\big\}c(X)\big] \\
&= k^*(\mathbf{E}[\tilde{g}(X)c(X)] - \mathbf{E}[g^*(X)c(X)]) \le 0,
\end{aligned}$$

where the first inequality follows from (5.5), and the second inequality uses the fact that $\tilde{g}$ satisfies the budget constraint and g^* meets the budget constraint exactly, and that $k^* > 0$ in this case. This implies

$$\mathbf{E}[\tilde{g}(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\tau(X)] \le \mathbf{E}[g^*(X)\mathbf{1}\{X \in \mathcal{X}_\epsilon\}\tau(X)],$$

so that g^* is optimal. ■

Lemma 5.1. *Under Assumption 2.5, for bounded positive ξ_n such that $n_S^{-\alpha} = o\big(\xi_n^2\big)$,*

$$\sup_{x:\tau(x)>\xi_n} \left|\delta(x;\hat{F}_1,\hat{F}_0) - \delta(x;F_1,F_0)\right| = O_p\big(\xi_n^{-2} n_S^{-\alpha}\big),$$

as $n_S \to \infty$.

Proof: Denote

$$\begin{aligned}
\Delta(y \mid x) &= F_1(y \mid x) - F_0(y \mid x), \text{ and} \\
\hat{\Delta}(y \mid x) &= \hat{F}_1(y \mid x) - \hat{F}_0(y \mid x).
\end{aligned}$$

We can write

$$\sup_{x:\tau(x)>\xi_n} \left|\delta(x;\hat{F}_1,\hat{F}_0) - \delta(x;F_1,F_0)\right|$$

$$
\begin{aligned}
&= \sup_{x:\tau(x)>\xi_n} \left| \frac{\int \hat{\Delta}(y\mid x)\mathbf{1}\left\{\hat{\Delta}(y\mid x)>0\right\}dy}{\int \left|\hat{\Delta}(y\mid x)\right|dy} - \frac{\int \Delta(y\mid x)\mathbf{1}\left\{\Delta(y\mid x)>0\right\}dy}{\int |\Delta(y\mid x)|dy} \right| \\
&\le \sup_{x:\tau(x)>\xi_n} \left| \frac{\int \hat{\Delta}(y\mid x)\mathbf{1}\left\{\hat{\Delta}(y\mid x)>0\right\}dy}{\int \left|\hat{\Delta}(y\mid x)\right|dy} - \frac{\int \hat{\Delta}(y\mid x)\mathbf{1}\left\{\hat{\Delta}(y\mid x)>0\right\}dy}{\int |\Delta(y\mid x)|dy} \right| \\
&\quad + \sup_{x:\tau(x)>\xi_n} \left| \frac{\int \hat{\Delta}(y\mid x)\mathbf{1}\left\{\hat{\Delta}(y\mid x)>0\right\} - \Delta(y\mid x)\mathbf{1}\left\{\Delta(y\mid x)>0\right\}dy}{\int |\Delta(y\mid x)|dy} \right| \\
&\le \sup_{x:\tau(x)>\xi_n} \int \left|\hat{\Delta}(y\mid x)\right| \left| \frac{1}{\int \left|\hat{\Delta}(y\mid x)\right|dy} - \frac{1}{\int |\Delta(y\mid x)|dy} \right| dy \\
&\quad + \sup_{x:\tau(x)>\xi_n} \frac{1}{\xi_n} \left| \int \left[\hat{\Delta}(y\mid x) - \Delta(y\mid x)\right]\mathbf{1}\left\{\hat{\Delta}(y\mid x)>0\right\}dy \right| \\
&\quad + \sup_{x:\tau(x)>\xi_n} \frac{1}{\xi_n} \left| \int \Delta(y\mid x)\left(\mathbf{1}\left\{\hat{\Delta}(y\mid x)>0\right\} - \mathbf{1}\left\{\Delta(y\mid x)>0\right\}\right)dy \right| \qquad (5.6)
\end{aligned}
$$

where second inequality is due to

$$
\xi_n < \tau(x) = \int_{-M}^{M} F_0(y\mid x) - F_1(y\mid x)dy \le \int_{-M}^{M} |\Delta(y\mid x)|dy.
$$

Since outcomes are bounded, the first term in (5.6) is bounded above by

$$
\begin{aligned}
&\sup_{x:\tau(x)>\xi_n} 2M \left| \frac{1}{\int \left|\hat{\Delta}(y\mid x)\right|dy} - \frac{1}{\int |\Delta(y\mid x)|dy} \right| \\
&\le \frac{2M}{\xi_n} \frac{\sup_{x:\tau(x)>\xi_n} \left| \int |\Delta(y\mid x)| - \left|\hat{\Delta}(y\mid x)\right| dy \right|}{\inf_{x:\tau(x)>\xi_n} \int \left|\hat{\Delta}(y\mid x)\right|dy} \\
&\le \frac{2M}{\xi_n} \frac{\int \sup_{x\in\mathcal{X},y\in\mathcal{Y}} \left|\hat{\Delta}(y\mid x) - \Delta(y\mid x)\right|dy}{\inf_{x:\tau(x)>\xi_n} \int |\Delta(y\mid x)|dy - \sup_{x:\tau(x)>\xi_n} \int \left|\hat{\Delta}(y\mid x) - \Delta(y\mid x)\right|dy} \\
&\le \frac{4M^2}{\xi_n} \frac{O_p(n_S^{-\alpha})}{\xi_n - O_p(n_S^{-\alpha})} = O_p\left(\xi_n^{-2} n_S^{-\alpha}\right).
\end{aligned}
$$

The second term in (5.6) is bounded by

$$
\sup_{x:\tau(x)>\xi_n} \frac{1}{\xi_n} \int \left|\hat{\Delta}(y\mid x) - \Delta(y\mid x)\right|dy \le \frac{2M}{\xi_n} \sup_{x,y:\tau(x)>\xi_n} \left|\hat{\Delta}(y\mid x) - \Delta(y\mid x)\right| = O_p\left(\xi_n^{-1} n_S^{-\alpha}\right).
$$

The last term in (5.6) is bounded by

$$\begin{aligned}
&\sup_{x:\tau(x)\geq\xi_n^{-1}}\int|\Delta(y\mid x)|\left|\mathbf{1}(\Delta(y\mid x)>0)-\mathbf{1}\big(\hat{\Delta}(y\mid x)>0\big)\right|dy\\
&=\sup_{x:\tau(x)\geq\xi_n^{-1}}\int|\Delta(y\mid x)|\mathbf{1}\big\{\Delta(y\mid x)-\hat{\Delta}(y\mid x)\geq\Delta(y\mid x)>0\big\}dy\\
&\quad+\sup_{x:\tau(x)\geq\xi_n^{-1}}\int|\Delta(y\mid x)|\mathbf{1}\big\{0\geq\Delta(y\mid x)>\Delta(y\mid x)-\hat{\Delta}(y\mid x)\big\}dy\\
&\leq 4M\sup_{x\in\mathcal{X},y\in\mathcal{Y}}\left|\Delta(y\mid x)-\hat{\Delta}(y\mid x)\right|dy=O_p(n_S^{-\alpha}).
\end{aligned}$$

Thus overall it is $O_p\big(\xi_n^{-2}n_S^{-\alpha}\big)$. ■

Lemma 5.2. *Under Assumption 2.5, for sequences $\epsilon_n\to\epsilon$, $\lambda_n\to 0$ and bounded positive ξ_n such that $n_S^\alpha\xi_n^2\lambda_n\to\infty$, the event*

$$\big\{x\in\mathcal{X}_{\epsilon_n-\lambda_n}:\tau(x)>\xi_n\big\}\subseteq\big\{x\in\hat{\mathcal{X}}_{\epsilon_n}:\tau(x)>\xi_n\big\}\subseteq\big\{x\in\mathcal{X}_{\epsilon_n+\lambda_n}:\tau(x)>\xi_n\big\}$$

occurs with probability approaching 1.

Proof: Denote the complement of a set S in the probability space by S^c. Since

$$\begin{aligned}
P((A\cap E)\subseteq(B\cap E)\subseteq(D\cap E))&=P(\{(A\cap E)\cap(B\cap E)^c\}\cup\{(B\cap E)\cap(D\cap E)^c\}=\varnothing)\\
&=P(\{(A\cap E)\cap(B^c\cup E^c)\}\cup\{(B\cap E)\cap(D^c\cup E^c)\}=\varnothing)\\
&=P((A\cap E\cap B^c)\cup(B\cap E\cap D^c)=\varnothing)\\
&=1-P((A\cap E\cap B^c)\cup(B\cap E\cap D^c)\neq\varnothing)\\
&\geq 1-P(A\cap E\cap B^c\neq\varnothing)-P(B\cap E\cap D^c\neq\varnothing),
\end{aligned}$$

we can write

$$\begin{aligned}
&P\big(\big\{x\in\mathcal{X}_{\epsilon_n-\lambda_n},\tau(x)>\xi_n\big\}\subseteq\big\{x\in\hat{\mathcal{X}}_{\epsilon_n},\tau(x)>\xi_n\big\}\subseteq\big\{x\in\mathcal{X}_{\epsilon_n+\lambda_n},\tau(x)>\xi_n\big\}\big)\\
&\geq 1-P\Big(\big\{x\in\mathcal{X}_{\epsilon_n-\lambda_n}\cap\hat{\mathcal{X}}_{\epsilon_n}^c,\tau(x)>\xi_n\big\}\neq\varnothing\Big)-P\Big(\big\{x\in\hat{\mathcal{X}}_{\epsilon_n}\cap\mathcal{X}_{\epsilon_n+\lambda_n}^c,\tau(x)>\xi_n\big\}\neq\varnothing\Big)\\
&=1-P\big\{\delta(x;F_1,F_0)\leq\epsilon_n-\lambda_n,\delta(x;\hat{F}_1,\hat{F}_0)>\epsilon_n\quad\text{for some } x \text{ with } \tau(x)>\xi_n\big\}\\
&\quad-P\big\{\delta(x;F_1,F_0)>\epsilon_n+\lambda_n,\delta(x;\hat{F}_1,\hat{F}_0)\leq\epsilon_n\quad\text{for some } x \text{ with } \tau(x)>\xi_n\big\}\\
&\geq 1-2P\left\{\sup_{x:\tau(x)>\xi_n}\left|\delta\big(x;\hat{F}_1,\hat{F}_0\big)-\delta(x;F_1,F_0)\right|>\lambda_n\right\}\to 1,\ \text{as } n_S\to\infty,
\end{aligned}$$

where the last line uses Lemma 5.1. ■

Lemma 5.3.

$$\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>k\right\}\hat{c}(X_i)-\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}\hat{c}(X_i)\right|\to_P 0,$$

as $n_T,n_S\to\infty$.

Proof: We write

$$\begin{aligned}
&P\left(\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\left(\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>k\right\}-\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}\right)\hat{c}(X_i)\right|>\epsilon\right)\\
&\leq P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\right\}|\mathbf{1}\{\rho(X_i)>k\}-\mathbf{1}\{\hat{\rho}(X_i)>k\}|>\frac{\epsilon}{2\overline{c}}\right)\\
&\quad+P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\left|\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\right\}-\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\right|\mathbf{1}\{\rho(X_i)>k\}>\frac{\epsilon}{2\overline{c}}\right).
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
&P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\right\}|\mathbf{1}\{\rho(X_i)>k\}-\mathbf{1}\{\hat{\rho}(X_i)>k\}|>\frac{\epsilon}{2\overline{c}}\right)\\
&=P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\left|\mathbf{1}\left(\rho(X_i)>k\geq\frac{\hat{\tau}(X_i)}{c(X_i)}\right)+\mathbf{1}(\hat{\rho}(X_i)>k\geq\rho(X_i))\right|>\frac{\epsilon}{2\overline{c}}\right)\\
&\leq P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left(\rho(X_i)\in\left[k,k+\frac{\xi_n}{2}\right]\right)>\frac{\epsilon}{4\overline{c}}\right)+P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left(\rho(X_i)\in\left[k-\frac{\xi_n}{2},k\right]\right)>\frac{\epsilon}{4\overline{c}}\right)\\
&\quad+P\left(\sup_{x\in\mathcal{X}}|\rho(x)-\hat{\rho}(x)|>\frac{\xi_n}{2}\right)\\
&\leq P\left(\sup_{k\geq\xi_n}(\mathbb{E}_n-\mathbb{E})\mathbf{1}\left(\rho(X_i)\in\left[k,k+\frac{\xi_n}{2}\right]\right)>\frac{\epsilon}{4\overline{c}}-\sup_{k\geq\xi_n}P\left(\rho(X_i)\in\left[k,k+\frac{\xi_n}{2}\right]\right)\right)\\
&+P\left(\sup_{k\geq\xi_n}(\mathbb{E}_n-\mathbb{E})\mathbf{1}\left(\rho(X_i)\in\left[k-\frac{\xi_n}{2},k\right]\right)>\frac{\epsilon}{4\overline{c}}-\sup_{k\geq\xi_n}P\left(\rho(X_i)\in\left[k-\frac{\xi_n}{2},k\right]\right)\right)\\
&\quad+o_P(1)
\end{aligned}$$

where $\mathbb{E}_n$ denotes the empirical mean operator, and the second inequality follows as $n_S^\alpha\xi_n\to\infty$. For given $\epsilon>0$, by bounded density, for large enough n, we have

$$\sup_{k\geq\xi_n}P\left(\rho(X_i)\in\left[k,k+\frac{\xi_n}{2}\right]\right)\leq\frac{\epsilon}{8\overline{c}},\text{ and}$$
$$\sup_{k\geq\xi_n}P\left(\rho(X_i)\in\left[k-\frac{\xi_n}{2},k\right]\right)\leq\frac{\epsilon}{8\overline{c}}.$$

Thus we have

$$
\begin{aligned}
&P\left(\sup_{k\geq\xi_n}(\mathbb{E}_n-\mathbb{E})\mathbf{1}\left(\rho(X_i)\in\left[k,k+\frac{\xi_n}{2}\right]\right)>\frac{\epsilon}{4\bar{c}}-\sup_{k\geq\xi_n}P\left(\rho(X_i)\in\left[k,k+\frac{\xi_n}{2}\right]\right)\right)\\
&+P\left(\sup_{k\geq\xi_n}(\mathbb{E}_n-\mathbb{E})\mathbf{1}\left(\rho(X_i)\in\left[k-\frac{\xi_n}{2},k\right]\right)>\frac{\epsilon}{4\bar{c}}-\sup_{k\geq\xi_n}P\left(\rho(X_i)\in\left[k-\frac{\xi_n}{2},k\right]\right)\right)\\
&\leq 2P\left(\sup_{B\in\mathcal{B}}(\mathbb{E}_n-\mathbb{E})\mathbf{1}(\rho(X_i)\in B)>\frac{\epsilon}{8\bar{c}}\right)
\end{aligned}
$$

where $\mathcal{B}$ is the set of closed balls in $\mathbf{R}$. By Kosorok (2008), Lemma 9.8 and 9.12, the collection of function $\mathbf{1}(\rho(X_i)\in B : B\in\mathcal{B})$ is VC subgraph and thus Glivenko-Cantelli, so this term goes to 0.

On the other hand, pick $\lambda_n\to 0$ monotonically such that $n_S^{\alpha}\xi_n^2\lambda_n\to\infty$. Then

$$
\begin{aligned}
&P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\left|\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\right\}-\mathbf{1}\{X_i\in\mathcal{X}_{\epsilon}\}\right|\mathbf{1}\{\rho(X_i)>k\}>\frac{\epsilon}{2\bar{c}}\right)\\
&\leq P\left(\sup_{k\geq\xi_n}\frac{1}{n_T}\sum_{i=1}^{n_T}\left|\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\right\}-\mathbf{1}\{X_i\in\mathcal{X}_{\epsilon}\}\right|\mathbf{1}(\rho(X_i)>\xi_n)>\frac{\epsilon}{2\bar{c}}\right)\\
&\leq P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}^c_{\epsilon_n}\cap\mathcal{X}_{\epsilon}\text{ and }\tau(X_i)>\xi_n\underline{c}\right\}>\frac{\epsilon}{2\bar{c}}\right)\\
&\quad+P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\cap\mathcal{X}^c_{\epsilon}\cap\mathcal{X}^c_{\epsilon_n+\lambda_n},\tau(X_i)>\xi_n\underline{c}\right\}>\frac{\epsilon}{2\bar{c}}\right)\\
&\quad+P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left\{X_i\in\hat{\mathcal{X}}_{\epsilon_n}\cap\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_n+\lambda_n},\tau(X_i)>\xi_n\underline{c}\right\}>\frac{\epsilon}{2\bar{c}}\right).
\end{aligned}
$$

By Lemma 5.2, the sets $\left\{x\in\hat{\mathcal{X}}^c_{\epsilon_n}\cap\mathcal{X}_{\epsilon}:\tau(x)>\xi_n\underline{c}\right\}$ and $\left\{x\in\hat{\mathcal{X}}_{\epsilon_n}\cap\mathcal{X}^c_{\epsilon_n+\lambda_n}:\tau(x)>\xi_n\underline{c}\right\}$ are empty with probability going to 1, so the first two term goes to 0. Secondly, since the sets $\mathcal{X}_{\epsilon_n+\lambda_n}$ are nested by monotonicity of $\epsilon_n+\lambda_n$, for there exists n_0 such that $n>n_0$ implies $P\left(X_i\in\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_n+\lambda_n}\right)<\frac{\epsilon}{4\bar{c}}$, and $\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_n+\lambda_n}\subseteq\mathcal{X}^c_{\epsilon_{n_0}}\cap\mathcal{X}_{\epsilon_{n_0}+\lambda_{n_0}}$. Thus

$$
\begin{aligned}
&P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left(X_i\in\hat{\mathcal{X}}_{\epsilon_n}\cap\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_n+\lambda_n}\right)\mathbf{1}\left(\tau(X_i)>k\underline{c}\right)>\frac{\epsilon}{2\bar{c}}\right)\\
&\leq P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\left(X_i\in\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_{n_0}+\lambda_{n_0}}\right)>\frac{\epsilon}{2\bar{c}}\right)\\
&\leq P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\left\{\mathbf{1}\left(X_i\in\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_{n_0}+\lambda_{n_0}}\right)-P\left(X_i\in\mathcal{X}^c_{\epsilon}\cap\mathcal{X}_{\epsilon_{n_0}+\lambda_{n_0}}\right)\right\}>\frac{\epsilon}{4\bar{c}}\right).
\end{aligned}
$$

The last probability vanishes by the weak law of large numbers. ■

Proof of Theorem 2.2: To get consistency of $\hat{k}$, we show that

$$\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\big\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>k\big\}\hat{c}(X_i)-\mathbf{E}[\mathbf{1}\{X_i\in\mathcal{X}_\epsilon,\rho(X_i)>k^*\}c(X_i)]\right|=o_P(1). \tag{5.7}$$

Once we have this, if budget constraint is not binding for setting cutoff at $k=0$, then for any $\eta>0$, we have

$$r_\eta:=B-\mathbf{E}[\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>\eta\}c(X_i)]>0$$

and

$$\begin{aligned}
P\big(\hat{k}>\eta\big)&=P\left(\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\big\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>\eta\big\}\hat{c}(X_i)>B\right)\\
&\leq P\left(\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\big\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>\eta\big\}\hat{c}(X_i)-\mathbf{E}[\mathbf{1}\{X_i\in\mathcal{X}_\epsilon,\rho(X_i)>\eta\}c(X_i)]\right|>r_\eta\right)\\
&\leq P\left(\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\big\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>k\big\}\hat{c}(X_i)-\mathbf{E}[\mathbf{1}\{X_i\in\mathcal{X}_\epsilon,\rho(X_i)>k\}c(X_i)]\right|>r_\eta\right)\\
&\to 0
\end{aligned}$$

where the second inequality holds for large enough n such that $\xi_n<\eta$. On the other hand, if budget constraint is binding, then consistency follows from standard arguments of extremum estimators, with identification condition given by the positive density of ρ around k^* under Assumption 2.5.

Decomposing LHS in (5.7),

$$\begin{aligned}
&\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\big\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>k\big\}\hat{c}(X_i)-\mathbf{E}[\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}c(X_i)]\right|\\
&\leq\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\big\{X_i\in\hat{\mathcal{X}}_{\epsilon_n},\hat{\rho}(X_i)>k\big\}\hat{c}(X_i)-\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}\hat{c}(X_i)\right|\\
&\quad+\sup_{k\geq\xi_n}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}\hat{c}(X_i)-\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}c(X_i)\right|\\
&\quad+\sup_{k\geq 0}\left|\frac{1}{n_T}\sum_{i=1}^{n_T}\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}c(X_i)-\mathbf{E}[\mathbf{1}\{X_i\in\mathcal{X}_\epsilon\}\mathbf{1}\{\rho(X_i)>k\}c(X_i)]\right|. \qquad (5.8)
\end{aligned}$$

The second term here is $o_P(1)$ by uniform convergence of $\hat{c}$. As for the third term, note that $g_k(\cdot)=\{\mathbf{1}(\cdot\in\mathcal{X}_\epsilon)\mathbf{1}(\rho(\cdot)>k)c(\cdot)\}$ is a bounded decreasing process in k. By Lemma 9.10 in

Kosorok (2008), it is VC subgraph with index 2, and is therefore Glivenko-Cantelli. This means the third line goes to 0 in probability. The first term is $o_P(1)$ by Lemma 5.3. ■

Lemma 5.4. *Under Assumption 2.5, let $\epsilon_n \to \epsilon$ be monotone decreasing, and bounded positive ξ_n be such that $n_S^\alpha \xi_n^2(\epsilon_n - \epsilon) \to \infty$. Then,*

$$\int \mathbf{1}\left\{x \in \mathcal{X} : \mathbf{1}\left\{\rho(x) > \xi_n, x \in \hat{\mathcal{X}}_{\epsilon_n}\right\} \neq \mathbf{1}\{\rho(x) > \xi_n, x \in \mathcal{X}_\epsilon\}\right\} dP_X(x) \to_P 0,$$

as $n_S \to \infty$.

Proof: The random subset is equal to the union of the following two random subsets:

$$\left\{x \in \mathcal{X}_\epsilon \cap \hat{\mathcal{X}}_{\epsilon_n}^c : \rho(x) > \xi_n\right\} \text{ and } \left\{x \in \mathcal{X}_\epsilon^c \cap \hat{\mathcal{X}}_{\epsilon_n} : \rho(x) > \xi_n\right\}.$$

We will show that the probability (push-forward) measures of these random subsets of $\mathcal{X}$ goes to 0 in probabiliy.

Observe that $\rho(x) > \xi_n$ implies $\tau(x) > \underline{c}\xi_n$ for some $\underline{c} > 0$ due to Assumption 2.5. By setting $\lambda_n = \epsilon_n - \epsilon$, Lemma 5.2 implies that the first set is empty with probability approaching 1. As for the second random set, let $\lambda_n \to 0$ monotonically be such that $n_S^\alpha \xi_n^2 \lambda_n \to \infty$. Then,

$$\begin{aligned}
&P_X\left\{x \in \mathcal{X}_\epsilon^c \cap \hat{\mathcal{X}}_{\epsilon_n} \text{ and } \rho(x) > \xi_n\right\} \\
&= P_X\left\{x \in \mathcal{X}_\epsilon^c \cap \hat{\mathcal{X}}_{\epsilon_n} \cap \mathcal{X}_{\epsilon_n+\lambda_n}^c \text{ and } \tau(x) > \xi_n \underline{c}\right\} + P_X\left\{x \in \mathcal{X}_\epsilon^c \cap \hat{\mathcal{X}}_{\epsilon_n} \cap \mathcal{X}_{\epsilon_n+\lambda_n} \text{ and } \tau(x) > \xi_n \underline{c}\right\} \\
&\leq P_X\left\{x \in \hat{\mathcal{X}}_{\epsilon_n} \cap \mathcal{X}_{\epsilon_n+\lambda_n}^c \text{ and } \tau(x) > \xi_n \underline{c}\right\} + P_X\left\{x \in \mathcal{X}_\epsilon^c \cap \mathcal{X}_{\epsilon_n+\lambda_n}\right\},
\end{aligned}$$

because $\epsilon_n + \lambda_n = 2\epsilon_n - \epsilon \to \epsilon$.

By Lemma 5.2, the first probability in the last line vanishes as $n_S \to \infty$. As for the second term, let $E_n = \mathcal{X}_\epsilon^c \cap \mathcal{X}_{\epsilon_n+\lambda_n}$. Then, $\bigcap_{n=1}^\infty E_n = \emptyset$. By construction, the sets are nested such that $E_1 \supseteq E_2 \supseteq E_3 \supseteq \ldots$. Furthermore, $P_X(E_1) \leq 1$. By continuity of measure,

$$\lim_{n\to\infty} P_X\{E_n\} = P_X\left\{\bigcap_{n=1}^\infty E_n\right\} = 0.$$

Hence,

$$P\left\{P_X\left\{x \in \mathcal{X}_\epsilon^c \cap \mathcal{X}_{\epsilon_n+\lambda_n}\right\} > \eta/2\right\} \to 0,$$

as $n_S \to \infty$. ■

Proof of Theorem 2.3: We can write

$$\begin{aligned}
&\left|\mathbf{1}\left\{x \in \hat{\mathcal{X}}_{\epsilon_n}, \hat{\rho}(x) > \hat{k}\right\} - \mathbf{1}\{x \in \mathcal{X}_\epsilon, \rho(x) > k^*\}\right| \\
&\leq \left|\mathbf{1}\{\rho(x) > k^*\} - \mathbf{1}\left\{\hat{\rho}(x) > \hat{k}\right\}\right| + \mathbf{1}\{\rho(x) > k^*\}\left|\mathbf{1}(x \in \mathcal{X}_\epsilon) - \mathbf{1}\left(x \in \hat{\mathcal{X}}_{\epsilon_n}\right)\right| \\
&\leq \left|\mathbf{1}\{\rho(x) > k^*\} - \mathbf{1}\left\{\hat{\rho}(x) > \hat{k}\right\}\right| + \mathbf{1}\{\rho(x) > \xi_n\}\left|\mathbf{1}(x \in \mathcal{X}_\epsilon) - \mathbf{1}\left(x \in \hat{\mathcal{X}}_{\epsilon_n}\right)\right|
\end{aligned}$$

$$+\mathbf{1}\{k^* < \rho(x) \le \xi_n\}.$$

For the second term, Lemma 5.4 states that

$$P_X\Big(\big\{x : \rho(x) > \xi_n, \mathbf{1}(x \in \mathcal{X}_\epsilon) \neq \mathbf{1}\big(x \in \hat{\mathcal{X}}_{\epsilon_n}\big)\big\}\Big) \to_p 0.$$

For the third term, if $k^* > 0$, then the set of x such that $\rho(x) \in (k^*, \xi_n]$ is empty for large enough n. On the other hand if $k^* = 0$, the probability measure goes to 0 due to bounded density for positive values of $\rho(X_i)$, so that

$$P_{X_i}(x : k^* < \rho(x) \le \xi_n) \to 0.$$

We tackle the first term by breaking into two cases. Suppose $k^* > 0$. Given $\epsilon > 0$, we can choose η_1, η_2 small enough such that if $\left|\hat{k} - k^*\right| \le \eta_1$ and $\sup_{x\in\mathcal{X}}|\hat{\rho}(x) - \rho(x)| \le \eta_2$ implies

$$\begin{aligned}
&P_{X_i}\Big(\big\{x : \mathbf{1}\{\rho(x) > k^*\} \neq \mathbf{1}\big\{\hat{\rho}(x) > \hat{k}\big\}\big\}\Big) \\
&\le P_{X_i}\Big(\big\{x : \hat{\rho}(x) > \hat{k}, \rho(x) \le k^*\big\}\Big) + P_{X_i}\Big(\big\{x : \hat{\rho}(x) \le \hat{k}, \rho(x) > k^*\big\}\Big) \\
&\le P_{X_i}(\{x : 0 < k^* - \eta_1 - \eta_2 < \rho(x) \le k^*\}) + P_{X_i}(\{k^* < \rho(x) \le k^* + \eta_1 + \eta_2\}) \\
&<\epsilon,
\end{aligned}$$

where the last line uses the bounded density of $\rho(X_i)$ for positive values. Assumption 2.5 and consistency of $\hat{k}$ imply that

$$\begin{aligned}
&P\Big(P_{X_i}\Big(\big\{x : \mathbf{1}\{\rho(x) > k^*\} \neq \mathbf{1}\big\{\hat{\rho}(x) > \hat{k}\big\}\big\}\Big) > \epsilon\Big) \\
&\le P\Big(\left|\hat{k} - k^*\right| > \eta_1\Big) + P\left(\sup_{x\in\mathcal{X}}|\hat{\rho}(x) - \rho(x)| > \eta_2\right) \to 0.
\end{aligned}$$

For the other case let $k^* = 0$. Given $\epsilon > 0$, we can pick $\eta_1 > 0$ such that on the event $\sup_{x\in\mathcal{X}}|\hat{\rho}(x) - \rho(x)| \le \xi_n$ and $\hat{k} \le \eta_1$, we have for large enough n,

$$\begin{aligned}
&P_{X_i}\Big(\big\{x : \mathbf{1}\{\rho(x) > k^*\} \neq \mathbf{1}\big\{\hat{\rho}(x) > \hat{k}\big\}\big\}\Big) \\
&\le P_{X_i}(\{x : \hat{\rho}(x) > \xi_n, \rho(x) \le 0\}) + P_{X_i}\Big(\big\{x : \hat{\rho}(x) \le \hat{k}, \rho(x) > 0\big\}\Big) \\
&\le 0 + P_{X_i}(\{x : 0 < \rho(x) \le \eta_1 + \xi_n\}) \le \epsilon,
\end{aligned}$$

where in the first inequality we have that $\hat{k} \ge \xi_n$ and the last inequality is again due to bounded density. And so

$$\begin{aligned}
&P\Big(P_{X_i}\Big(\big\{x : \mathbf{1}\{\rho(x) > k^*\} \neq \mathbf{1}\big\{\hat{\rho}(x) > \hat{k}\big\}\big\}\Big) > \epsilon\Big) \\
&\le P\left(\sup_{x\in\mathcal{X}}|\hat{\rho}(x) - \rho(x)| > \xi_n\right) + P\Big(\hat{k} \le \eta_1\Big) \to 0.
\end{aligned}$$

■

References

ALCANTUD, J. C. R., M. MARIOTTI, AND R. VENEZIANI (2022): *Theoretical Economics*(17), 1529–1557.

ALLCOTT, H. (2015): "Site Selection Bias in Program Evaluation," *Quarterly Journal of Economics*, 130(3), 1117–1165.

ATHEY, S., AND S. WAGER (2021): *Econometrica*(89), 133–161.

BHATTACHARYA, D., AND P. DUPAS (2012): "Inferring Welfare Maximizing Treatment Assignment Under Budget Constraints," *Journal of Econometrics*, 167, 168–196.

BOSSERT, W., S. CATO, AND K. KAMAGA (2023): *Economic Theory*(75), 1099–1139.

CHERNOZHUKOV, V., S. LEE, A. ROSEN, AND L. SUN (2025): "Policy Learning with Confidence," *arXiv:2502.10653v1 [econ.EM]*.

DUFLO, E. (2004): "Scaling Up and Evaluation," *Annual World Bank Conference on Development Economics*, pp. 341–369.

GECHTER, M. (2024): "Generalizing the Results from Social Experiments: Theory and Evidence from India," *Journal of Business and Economic Statistics*, 42, 801–811.

GECHTER, M., AND R. MEAGER (2022): "Combining Experimental and Observational Studies in Meta-Analysis: A Debiasing Approach," *Working Paper*.

HOTZ, J. V., G. W. IMBENS, AND J. H. MORTIMER (2005): "Predicting the Efficacy of Future Training Programs Using Past Experiences at Other Locations," *Journal of Econometrics*, 125, 241–270.

KITAGAWA, T., AND A. TETENOV (2018): "Who Should be Treated? Empirical Welfare Maximization Methods for Treatment Choice," *Econometrica*, 86, 591–616.

KOSOROK, M. R. (2008): *Introduction to Empirical Processes and Semiparametric Inference*. Springer, New York.

LESHNO, M., AND H. LEVY (2002): "Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance," *Management Science*, 48(48), 1074–1085.

MANSKI, C. F. (2004): "Statistical Treatment Rules for Heterogeneous Populations," *Econometrica*, 72(4), 1221–1246.

MURALIDHARAN, K., AND P. NIEHAUS (2017): "Experimentation at Scale," *Journal of Economic Perspectives*, 31, 103–214.

SUN, H., E. MUNRO, G. KALASHNOV, S. DU, AND S. WAGER (2025): "Treatment Allocation under Uncertain Costs," *arXiv.2103.11066*.

SUN, L. (2025): *Working Paper*.

VIVIANO, D., AND J. BRADIC (2024): *Journal of the American Statistical Association*(119), 730–743.

WANG, S., AND D. Y. YANG (2021): "Policy Experimentation in China: The Political Economy of Policy Learning," Discussion paper, National Bureau of Economic Research.