Policy-Aligned Estimation of Conditional Average Treatment Effects*

Artem Timoshenko[†] and Caio Waisman[†]

[†]Kellogg School of Management, Northwestern University

December 15, 2025

Abstract

Firms often develop targeting policies to personalize marketing actions and improve incremental profits. Effective targeting depends on accurately separating customers with positive versus negative treatment effects. We propose an approach to estimate the conditional average treatment effects (CATEs) of marketing actions that aligns their estimation with the firm's profit objective. The method recognizes that, for many customers, treatment effects are so extreme that additional accuracy is unlikely to change the recommended actions. However, accuracy matters near the decision boundary, as small errors can alter targeting decisions. By modifying the firm's objective function in the standard profit maximization problem, our method yields a near-optimal targeting policy while simultaneously estimating CATEs. This introduces a new perspective on CATE estimation, reframing it as a problem of profit optimization rather than prediction accuracy. We establish the theoretical properties of the proposed method and demonstrate its performance and trade-offs using synthetic data.

Keywords: Targeted Marketing; Uplift Modeling; Policy Learning; Conditional Average Treatment Effects; Causal Inference; M-estimation.

*We would like to thank audiences at the 2025 Conference on Digital Experimentation, the 2025 Conference on Artificial Intelligence, Machine Learning, and Business Analytics, and the Kellogg Ad-Tech Lab. Glenn Feng provided outstanding research assistance. Contact: artem.timoshenko@northwestern.edu (AT) and caio.waisman@northwestern.edu (CW)

1 Introduction

Firms increasingly personalize marketing actions to improve incremental profits. Retailers decide which customers should receive discounts, car dealers personalize pricing, and online platforms target sponsored ads (Simester et al., 2020b; Dubé and Misra, 2023; Rafieian and Yoganarasimhan, 2021). These decisions rely on targeting models that recommend the most effective marketing action for each customer.

There are two conceptually distinct approaches to developing targeting models. The first estimates the impact of marketing actions conditional on customer characteristics (conditional average treatment effects; CATEs) and recommends for each customer the action with the largest estimated treatment effect. The second approach bypasses CATE estimation; it directly learns a mapping from characteristics to actions by maximizing the expected policy value, such as the profit per customer. We refer to these approaches as "plug-in estimation" and "direct policy optimization," respectively.

Asymptotically, these two approaches are equivalent, but in finite samples they often lead to different targeting policies (Fernández-Loría and Provost, 2022). Direct policy optimization is appealing because it explicitly focuses on profit maximization, which is the ultimate goal for most firms. In contrast, plug-in estimation seeks accurate CATE estimates across the entire covariate space. When data are limited, attempting to approximate CATEs uniformly may reduce effectiveness near the decision boundary, where customers' treatment effects are close to the treatment cost. Nevertheless, plug-in methods offer transparency and interpretability, as CATEs can justify targeting decisions, support broader organizational adoption, and help implement fairness or volume constraints (Athey and Imbens, 2016; Lu et al., 2025).

We propose a framework to reconcile these two paradigms by reframing CATE estimation as an optimization problem. Our approach recognizes that the profit maximization objective in direct policy optimization focuses only on optimal binary decisions and does not identify the magnitude of treatment effects. We modify the profit maximization problem by replacing the known deterministic treatment costs with a stochastic threshold. This allows the resulting optimization problem to simultaneously recover a near-optimal targeting policy and estimate CATEs that rationalize this policy, which we refer to as "policy-aligned CATEs."

Our framework provides fine-grained control over the balance between policy optimization and CATE estimation. The distribution of a stochastic threshold governs this alignment. In the extremes, when the distribution is fully concentrated at the treatment cost, our optimization problem coincides with direct policy optimization. Conversely, under a uniform distribution, the problem is equivalent to standard plug-in estimation. By varying the density around treatment costs, our method modulates the trade-off between global CATE accuracy and precision around the decision boundary, thus optimizing finite-sample profits.

We prove the theoretical properties of the proposed approach and demonstrate them using synthetic data. The proposed surrogate objective function is Fisher consistent with respect to the original profit function, meaning that both share the same maximizer. This ensures that our estimator remains asymptotically valid for profit optimization. Furthermore, we show that our CATE estimator is an M-estimator. This implies statistical consistency under standard regularity conditions and allows for inference using conventional tools, even when the model includes many covariates or flexible functional forms.

Our contribution can be viewed from two complementary perspectives. First, from the policy learning perspective, we develop an estimator that yields near-optimal targeting policies while providing interpretable, policy-aligned CATEs. Our approach replaces the step-wise profit function in direct policy optimization with a smooth surrogate objective, a controllable deviation that enables CATE identification. These CATEs explain targeting decisions, facilitate counterfactual analyses such as last-minute circulation changes, and can serve as inputs in downstream optimization.

Second, we contribute to the CATE estimation literature by incorporating policy value as the primary estimation criterion. Traditional approaches estimate CATEs by minimizing global discrepancy measures, such as mean squared error (MSE). Because individual treatment effects are unobserved, researchers cannot validate CATEs directly. Instead, they typically evaluate the targeting policies implied by the CATE estimates using out-of-sample policy value. We depart from this convention by recasting the policy value from a validation criterion to an estimation criterion. Our estimator optimizes policy value directly while incorporating an explicit trade-off between precision near the decision boundary and global accuracy.

Beyond marketing and profit maximization, the proposed framework broadly applies to uplift-based targeting settings with binary treatments. This includes personalized medicine and adaptive clinical trials (Dahabreh et al., 2016), heterogeneous treatment assignment in economics (Dehejia, 2005), and resource allocation under uncertainty in operations research (Zhou et al., 2023). Implementation code is available at <link hidden for peer review>.

The rest of the paper proceeds as follows. First, we review the related literature on estimation of CATEs and policy learning. Section 2 presents our framework and outlines the assumptions. Section 3 introduces our proposed methodology and demonstrates that our surrogate objective function is Fisher consistent with respect to the firm's original profit function. Section 4 describes practical implementation through M-estimation. Section 5 documents the performance of our method using synthetic data. Section 6 concludes.

Related Literature

Our work relates to the literature on estimation of CATEs and on policy learning. We organize the discussion around these two streams and explain how our paper connects them.

Estimation of CATEs

The literature on CATE estimation is vast. Previous research used linear and kernel regressions, matching, and propensity score methods as reviewed by Imbens (2004). More recently, due to the availability of richer data, newer methods based on machine learning have been introduced to estimate CATEs; Jacob (2021) and Hu (2023) are recent reviews of these methods.

There are two important aspects of the CATE estimation literature that deserve attention in the context of our paper. First, using the terminology from Künzel et al. (2019), there are generally two types of methods for estimating CATEs. The first, referred to as T-learners, consists of fitting two models to estimate the conditional expectation of the outcome variable on covariates using treatment and control observations separately. The second, referred to as S-learners, employs a single model for estimation using all observations concurrently. Under this classification, our proposed method is an S-learner that allows CATE parameterization with flexible functional forms, including neural networks.

Second, most methods rely on the mean squared error (MSE) measure or its variants created by, for instance, adding a penalization term for regularization. This route is taken, for example, by Imai and Ratkovic (2013), Athey and Imbens (2016), Wager and Athey (2018), Nie and Wager (2021), and Hitsch et al. (2024). In contrast, our estimation criterion is derived directly from the decision maker's true optimization objective, which in our setting is the firm's expected profit.[1]

Three papers are most closely related to our approach. Elmachtoub and Grigas (2022) study linear optimization problems with unknown parameters, such as shortest path and portfolio optimization. They propose a "smart predict-then-optimize" framework that integrates the decision maker's payoff optimization with parameter prediction. To ensure computational tractability, Elmachtoub and Grigas (2022) introduce surrogate formulations to the original problem and prove their Fisher consistency, which means that the surrogate and the original loss share the same optimal solutions. In a similar spirit, our paper introduces a Fisher consistent surrogate objective for profit maximization that also produces consistent estimates of the CATEs. While Elmachtoub and Grigas (2022) focus on linear cost minimization tasks, we study an uplift-based targeting problem and causal estimation, in which the objective is the firm's expected incremental profit.

In turn, Ascarza and Israeli (2022) propose an algorithm for treatment effect estimation that modifies random forests to extract heterogeneity from features unrelated to protected attributes, thereby improving fairness in resource allocation. Similarly, our paper adjusts the estimation objective to emphasize the heterogeneity most relevant to the decision context. However, rather than focusing on fairness in the covariate space, we target heterogeneity in the profit outcomes. Our CATE estimator is most accurate near the decision boundary, where customers' treatment effects are close to treatment costs and small estimation errors can change targeting decisions.

Finally, in concurrent research, Chen et al. (2025) also aim to obtain more accurate CATE estimates around the decision boundary, but they focus on the experiment design rather than the estimation approach. Chen et al. (2025) consider a firm that has access to a pool of customers and can sequentially add them to the experiment. The authors propose a sampling procedure that uses customer characteristics to adaptively increase the sample size for CATE estimation via causal forests. In contrast, we take the experimental design and the corresponding propensity scores as given and obtain improved decision-making by altering the estimation procedure.

[1] The insight that MSE can be inadequate for firm decision-making dates back at least to Blattberg and George (1992), who studied price-elasticity estimation.

Policy Learning

The policy learning literature focuses on designing targeting policies that maximize incremental outcomes. A natural approach to policy learning is to first estimate the CATEs and then use these estimates to define the policy (Manski, 2004; Lemmens and Gupta, 2020). Several studies have derived the resulting properties of such policies; we refer the reader to Hirano and Porter (2020) for a comprehensive survey.

A different route circumvents the estimation of CATEs and instead recovers the optimal policy directly. This is possible by recognizing that the treatment assignment problem is closely related to a weighted classification problem, an insight that has been independently discovered in statistics (Zhao et al., 2012), computer science (Swaminathan and Joachims, 2015), and economics (Kitagawa and Tetenov, 2018). Recent research has extended direct policy learning to incorporate doubly robust estimators (Athey and Wager, 2021), model selection and penalization (Mbakop and Tabord-Meehan, 2021), and deep learning (Zhang, 2025). The distinction also resembles an ancient difference between "value-function" and "policy-value" optimization in dynamic programming.

Our proposed method is motivated by direct policy learning approaches. We modify the policy value optimization objective to obtain a near-optimal targeting policy together with policy-aligned CATE estimates.

2 Framework

We now introduce the firm's decision problem and the estimation assumptions. Although our primary motivation and empirical applications are in the marketing domain, the proposed approach is more general and can be applied for uplift-based resource allocation in different contexts. For ease of exposition, we will describe our framework with a focus on a firm that allocates marketing actions to maximize expected profit.

2.1 Problem Formulation

We consider a firm with a customer base in which each individual customer i is characterized by observable attributes, $X_i \in \mathbb{X} \subseteq \mathbb{R}^k$. These attributes can include demographic data and past purchasing histories. The firm assigns a binary treatment $W_i \in \{0, 1\}$ to each customer i. The treatment can represent various marketing actions, such as the allocation of coupons or catalog mailing decisions.

We denote the potential outcomes with and without treatment by $Y_i(1)$ and $Y_i(0)$, respectively. The potential outcomes can capture different target values for the firm. For example, the outcomes could represent profit per customer, store visits, or customer lifetime value projections (CLV). For simplicity, we assume that they are measured in monetary units. We define the conditional average treatment effect (CATE) for customers with characteristics X_i as

$$\tau_0(X_i) \equiv \mathbb{E}\left[Y_i(1) - Y_i(0) | X_i\right]. \tag{1}$$

The firm designs a targeting policy that assigns potentially different treatments to different customers based on observable characteristics, $\pi : \mathbb{X} \to \{0, 1\}$. The goal is to construct a targeting policy that maximizes the expected profit per customer ("policy value"):

$$\begin{aligned}
&\max_{\pi:\mathbb{X}\to\{0,1\}} \mathbb{E}_{Y(1),Y(0),X}\left[\pi(X_i)\left\{Y_i(1) - c\right\} + \left\{1 - \pi(X_i)\right\} Y_i(0)\right] \\
&\max_{\pi:\mathbb{X}\to\{0,1\}} \mathbb{E}_{X}\left[\pi(X_i)\left\{\tau_0(X_i) - c\right\}\right] \qquad (2)
\end{aligned}$$

where c is a known treatment cost per customer in the treatment condition, $W_i = 1$, and the expectation integrates over $Y_i(1)$, $Y_i(0)$ and X_i. We introduce c for expositional reasons. In many applications, this cost can be embedded in the definitions of potential outcomes, which is equivalent to $c = 0$ in our formulation.

If the CATEs were known to the firm, the optimal targeting policy, $\pi^*(X_i)$, would assign treatment only to customers with $\tau_0(X_i) \geq c$, that is,

$$\pi^*(X_i) = \mathbb{1}\left\{\tau_0(X_i) \geq c\right\}. \tag{3}$$

However, in practice, CATEs are unknown and must be estimated from data.

2.2 Data and Assumptions

We assume that the firm estimates a targeting policy using experimental data. In particular, we consider an experiment with n customers randomly assigned to treatment ($W_i = 1$) and control conditions ($W_i = 0$) with propensity scores e_i defined as

$$e_i = e(X_i) \equiv \Pr\left(W_i = 1 | X_i\right). \tag{4}$$

The firm's data include the experimental assignments, W_i, customer characteristics, X_i, propensity scores, e_i, and observed outcomes, Y_i, where $Y_i = W_i Y_i(1) + (1 - W_i) Y_i(0)$. We summarize our notation in Table 1.

Table 1: Notation

Notation	Definition
n	Number of customers in the experiment.
Y_i	Observed outcome for customer i in the experiment.
W_i	Treatment assignment for customer i in the experiment (binary).
X_i	Descriptive variables for customer i, $X_i \in \mathbb{R}^k$.
$e(X_i)$	Probability of treatment for customer i in the experiment, $\Pr(W_i = 1 \vert X_i)$.
c	Treatment cost per customer in the treatment condition ($W_i = 1$).
$\tau_0(X_i)$	Conditional average treatment effect, $\mathbb{E}[Y_i(1) - Y_i(0) \vert X_i]$.
$\pi(X_i)$	Policy recommendation for customer i, $\pi : \mathbb{X} \to \{0, 1\}$.

Throughout our analysis, we maintain the following assumptions.

Assumption 1. *Stable Unit Treatment Value Assumption (SUTVA):* $Y_i(W_1, \ldots, W_i, \ldots, W_n) = Y_i(W_i)$ *for all* i.

Assumption 2. *Unconfoundedness:* $Y_i(1), Y_i(0) \perp\!\!\!\perp W_i \vert X_i$ *for all* i.

Assumption 3. *Overlap:* $e(X) \in (0, 1)$ *for all* $X \in \mathbb{X}$.

Assumptions 1–3 frequently appear in the causal inference literature and are often plausible in marketing experiments. SUTVA ensures that there is no interference between customers in the experiment, so that the potential outcomes of any customer are not affected by the treatment assignments of other customers. Unconfoundedness and overlap are requirements on the experimental design. They ensure that treatment is not deterministic for certain values of X_i and, conditional on X_i, that treatment assignments are independent of potential outcomes.

2.3 Policy Training and Evaluation

There are two conceptually different approaches to estimate optimal targeting policies using experimental data. The first approach proceeds in two stages. It first obtains estimates of the CATEs, $\widehat{\tau}(\cdot)$, and then compares these estimates with the treatment cost c to obtain a targeting policy $\widehat{\pi}(\cdot) = \mathbb{1}\{\widehat{\tau}(\cdot) \geq c\}$. The CATE estimates are typically obtained by minimizing a measure of the discrepancy between the predicted treatment effects and the observed data. For example, we can parameterize the CATEs with a vector, θ, and compute it by minimizing the following mean squared error (MSE)

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \left[Y_i^* - \tau(X_i;\theta)\right]^2 \tag{5}$$

where Y_i^* is an unbiased estimate of the CATE of customer i, often referred to as a transformed outcome (Athey and Imbens, 2016), and is given by

$$Y_i^* = Y_i \left(\frac{W_i}{e(X_i)} - \frac{1 - W_i}{1 - e(X_i)} \right). \tag{6}$$

The second approach directly estimates the optimal targeting policy without the intermediate step of first estimating the CATEs. In particular, firms can parametrize the targeting policy $\pi(X_i;\theta)$ directly, construct a sample analog of Equation (2), and compute parameters that maximize the estimated policy value

$$\begin{aligned}
&\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \left[\frac{W_i \cdot \pi(X_i;\theta)}{e(X_i)} Y_i(1) - \frac{(1-W_i)[1-\pi(X_i;\theta)]}{1-e(X_i)} Y_i(0) - \pi(X_i;\theta)c \right] \\
&\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \left\{ \pi(X_i;\theta) \left[Y_i \left(\frac{W_i}{e(X_i)} - \frac{1-W_i}{1-e(X_i)} \right) - c \right] \right\} \\
&\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \left[\pi(X_i;\theta) \left(Y_i^* - c \right) \right] \qquad (7)
\end{aligned}$$

Intuitively, under Assumptions (1)–(3), Equation (7) provides an unbiased estimate of the policy value, as if the policy $\pi(X_i;\theta)$ were evaluated in a field experiment (Simester et al., 2020a; Hitsch et al., 2024). Firms can then directly search for the parameters θ that yield the targeting policy with the highest policy value.

Fernández-Loría and Provost (2022) recognize that these two estimation approaches are not equivalent. In finite samples, these approaches often yield targeting policies that are substantially different in recommended actions and out-of-sample policy value. Fernández-Loría and Provost (2022) argue that the bias-variance trade-off implied by fitting the CATEs is not necessarily aligned with that of the profit-maximizing objective, which can undermine policy learning via CATE estimation. Direct policy optimization methods do not suffer from this problem because they calibrate the policy parameters using the "correct" objective function.

The theoretical appeal of direct policy optimization comes at the expense of not estimating the CATEs, which are helpful in many business and research applications. CATEs explain why different customers are treated differently, which is useful for diagnostics during policy training and provides explainability for broader adoption. In business practice, CATE estimates can work together with optimization when implementing guardrails, such as volume or fairness constraints (Lu et al., 2025).

3 Proposed Method

Our proposed method modifies the profit maximization objective in Equation (7) to infer a near-optimal targeting policy and policy-aligned CATEs within the same optimization framework. To provide intuition about our methodology, we first describe it in a simpler context without covariates and using population moment conditions. We demonstrate that our surrogate objective function is Fisher consistent with respect to the original profit function, which justifies its use. Then, we incorporate covariates and argue that Fisher consistency continues to hold.

3.1 Intuition with No Covariates

Consider a hypothetical treatment allocation problem for an observationally equivalent group of customers ($X_i = X$ for all i), with the average treatment effect (ATE) given by $\tau_0 \equiv \mathbb{E}[Y_i(1) - Y_i(0)]$. The firm decides whether to allocate treatment to these customers, and akin to Equation (2), we can write the following optimization problem for the firm

$$
\begin{aligned}
&\max_{\pi \in \{0,1\}} \mathbb{E}_{Y(1),Y(0)} \left[\pi \left\{Y_i(1) - c\right\} + (1-\pi) Y_i(0)\right] \\
&\max_{\pi \in \{0,1\}} \pi \cdot (\tau_0 - c) \qquad (8)
\end{aligned}
$$

Can we use this profit maximization formulation to recover the ATE? For any treatment effect estimate τ, the recommended policy is $\pi = \mathbb{1}\left\{\tau \geq c\right\}$, so we can rewrite the optimization objective as follows:

$$
\max_{\tau \in \mathbb{R}} \mathbb{1}\left\{\tau \geq c\right\} (\tau_0 - c) \qquad (9)
$$

An important observation is that τ_0 is not a unique solution to the optimization problem (9). Indeed, this optimization problem implies a set of optimal solutions

$$\tilde{\tau} = \begin{cases} [c, +\infty) & \text{if } \tau_0 \geq c \\ (-\infty, c) & \text{otherwise.} \end{cases} \tag{10}$$

The true ATE τ_0 is one of the solutions to the profit maximization problem, but there are infinitely many optimal solutions. This is because optimal treatment allocation depends only on whether the true ATE exceeds the treatment cost, $\tau_0 \geq c$, rather than on the precise value of τ_0.

We propose to modify the objective function in Equation (9) so that the optimization problem is both close to the true profit optimization formulation and has a unique optimal solution at τ_0. Specifically, we estimate the ATE by optimizing the following surrogate objective function:

$$\hat{\tau} = \arg\max_{\tau \in \mathbb{R}} \mathbb{E}_C \left[\mathbb{1}\{\tau \geq C\} (\tau_0 - C) \right] \tag{11}$$

$$= \arg\max_{\tau \in \mathbb{R}} \int_{-\infty}^{\tau} (\tau_0 - u) f_C(u) du. \tag{12}$$

where we replace the (deterministic) treatment cost c in Equation (9) with a random variable C, drawn from a distribution of our choice, $F_C(\cdot)$ with density $f_C(\cdot)$.

Intuitively, the standard profit optimization function "rewards" estimates to make the right binary decisions; for example, if $\tau_0 > c$, then for any choice of $\tau > c$, the reward is $\tau_0 - c$, irrespective of the magnitude of τ. This means that τ can largely deviate from τ_0, and the estimator does not "learn" about it from the optimization loss. In contrast, the surrogate objective explicitly penalizes deviations from τ_0. The key insight is to view the treatment cost c as stochastic. Continuing the example, if $\tau > \tau_0$, then for realizations $C \in (\tau_0, \tau]$, the reward is strictly negative, and integration over C yields worse outcomes compared to choosing exactly τ_0. Similarly, for any $\tau < \tau_0$, the reward can be increased by raising τ. Thus, under a random cost C, the optimizer's expected reward is maximized by choosing $\tau = \tau_0$.[2]

[2] The proposed surrogate optimization problem is inspired by the literature on advertising auctions. Online advertisers often make decisions based on the distribution of competing bids, and for a given bid, the advertiser's outcomes are stochastic rather than deterministic. Waisman et al. (2025) leverage this insight to infer the effectiveness of online advertising. Conceptually, our approach requires a targeting policy to "submit bids" by choosing τ, and customers are only treated if this "bid" exceeds the (stochastic) treatment cost C, which plays the role of the highest competing bid. Similar to Waisman et al. (2025), the optimal bid coincides with the true ATE, τ_0.

Formally, the surrogate objective function is Fisher consistent with respect to the original objective function in Equation (9); in other words, $\hat{\tau}$ is also a maximizer of the original objective function (Lin, 2004). Furthermore, with a large sample, it follows that $\hat{\tau} = \tau_0$, so optimizing the surrogate function and measuring τ_0 are equivalent. Proposition 1 states these results, and the proof is provided in Appendix E.

Proposition 1. *If the density $f_C(\cdot)$ is strictly positive in a neighborhood of τ_0, then $\hat{\tau} = \tau_0$.*

The distribution $F_C(\cdot)$ controls the deviation of the proposed surrogate objective from the standard profit optimization. In the limit where the distribution is fully concentrated at c, the proposed optimization problem coincides with the standard profit optimization in Equation (9). In contrast, we demonstrate in Section 4.1 that for a uniform distribution $F_C(\cdot)$, the surrogate objective function is equivalent to minimizing MSE loss uniformly across the covariate space, as in Equation (6).

In Appendix A, we show that Equation (12) is equivalent to the following formulation

$$\hat{\tau} = \arg\max_{\tau \in \mathbb{R}} F_C(\tau) \left[\tau_0 - \kappa_C(\tau)\right], \tag{13}$$

where $\kappa_C(\tau) \equiv \mathbb{E}\left[C | C \leq \tau\right]$. This expression is especially convenient for the implementation of our method. We choose the distribution $F_C(\cdot)$ to ensure that it is centered around c, has a strictly positive density around τ_0, and the conditional expectation $\kappa_C(\cdot)$ is tractable. For example, in Section 4, we focus on the normal and logistic distributions.

Figure 1 illustrates the difference between the standard profit maximization problem and the proposed surrogate. On the horizontal axis, we plot different values of τ, and the vertical axis indicates the value of the optimization objective. For the optimization problem (9), the objective function is a stepwise function. For the modified optimization problem (11), the objective function is smooth and has a unique maximum at τ_0. The discrepancy between the two optimization objectives depends on the choice of the distribution $F_C(\cdot)$.[3]

3.2 Incorporating Covariates

We now extend the proposed estimation approach from a single homogeneous customer segment to the heterogeneous population with customer covariates. Instead of estimating the ATE τ_0, we aim to recover the CATE function $\tau_0 : \mathbb{X} \to \mathbb{R}$ using the profit optimization

[3] Figure 1 shows the surrogate objective function for $C \sim N(c, \sigma^2)$, and in Appendix B, we illustrate the objective function with different values of σ.

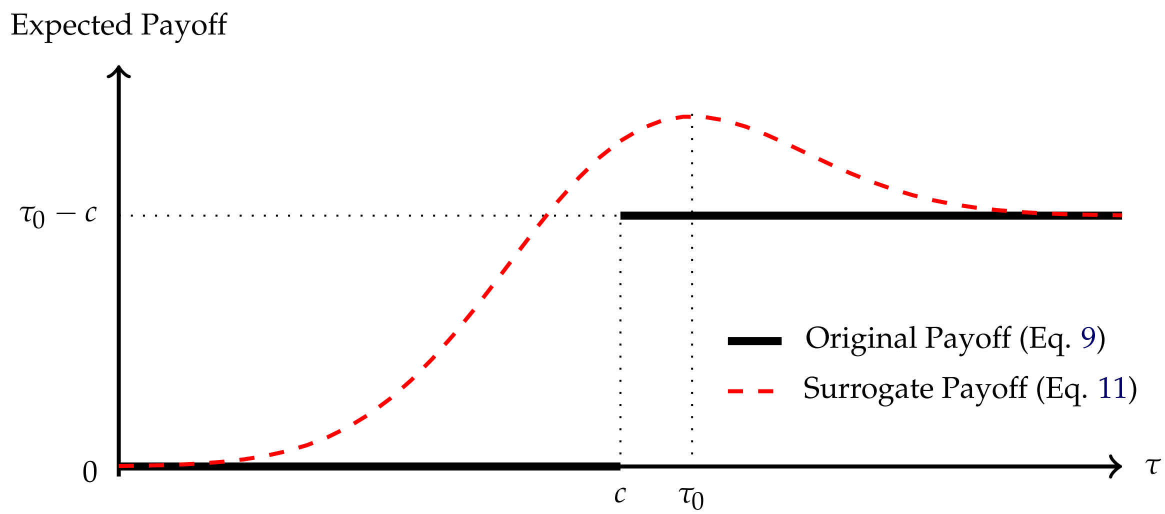

Figure 1: Illustration of the Surrogate Objective Function

objective:

$$\max_{\tau:\mathbb{X}\to\mathbb{R}} \mathbb{E}_X \left[\mathbb{1}\left\{\tau(X) \geq c\right\} \left\{\tau_0(X) - c\right\}\right] \tag{14}$$

We denote the (original) profit optimization objective as $Q(\tau)$:

$$Q(\tau) \equiv \mathbb{E}_X \left[\mathbb{1}\left\{\tau(X) \geq c\right\} \left\{\tau_0(X) - c\right\}\right]. \tag{15}$$

Similar to Equations (11) and (13), we propose the following surrogate objective:

$$\tilde{Q}(\tau) \equiv \mathbb{E}_{C,X} \left[\mathbb{1}\left\{\tau(X) \geq C\right\} \left\{\tau_0(X) - C\right\}\right] = \mathbb{E}_X \left(F_C\left[\tau(X)\right] \left\{\tau_0(X) - \kappa_C\left[\tau(X)\right]\right\}\right), \tag{16}$$

where for simplicity we assume that $C \perp\!\!\!\perp X$.

Proposition 2 establishes Fisher consistency for the surrogate objective function with covariates. The proof is provided in Appendix E.

Proposition 2. *Assume that the density $f_C(\cdot)$ is strictly positive on a neighborhood of $\tau_0(X)$ for all $X \in \mathbb{X}$. Then $\tilde{Q}(\tau)$ is maximized at $\tau_0(X)$ and is Fisher consistent with respect to $Q(\tau)$.*

3.3 Discussion

The proposed surrogate framework bridges the direct policy optimization and plug-in estimation approaches for targeting policies. It yields a Fisher-consistent estimator that identifies CATEs while remaining aligned with the firm's profit objective.One way to view the proposed estimator is that it emphasizes precision near the decision boundary. From a profit-optimization perspective, it is most valuable for firms to accurately estimate treatment effects for customers whose expected effects are close to the treatment cost. In contrast, for customers with very large or very small treatment effects, minor estimation errors are unlikely to alter targeting decisions.

We can compare our proposed method with an alternative approach that directly combines profit optimization and estimation accuracy for treatment effects. Specifically, similar to Equation (9), one could consider the following objective:

$$\max_{\tau \in \mathbb{R}} \mathbb{1}\{\tau \geq c\}(\tau_0 - c) + \lambda(\tau_0 - \tau)^2. \tag{17}$$

This objective blends the profit-maximization and CATE-fitting components, with the parameter λ controlling the trade-off between them. However, this formulation is discontinuous at $\tau = c$, and when extended to CATE estimation with covariates, it will exhibit discontinuities throughout the optimization space. In contrast, with a suitable choice of the surrogate distribution $F_C(\cdot)$, our proposed objective is continuously differentiable and well-behaved. By replacing the discontinuous indicator with a smoothed surrogate, our framework achieves the same decision-theoretic alignment while ensuring differentiability, enabling the use of gradient-based optimization methods.

4 Practical Implementation: M-Estimation

So far, we have presented the ideas behind our approach using population moments. In practice, we rely on a data sample to approximate these quantities. We estimate the near-optimal targeting policy and the policy-aligned CATEs by constructing a sample analog and incorporating covariates into Equation (13). Conceptually, our estimator is given by

$$\widehat{\tau}(\cdot) = \arg\max_{\tau:\mathbb{X}\to\mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} F_C\left[\tau(X_i)\right] \left\{ Y_i^* - \kappa_C\left[\tau(X_i)\right] \right\}. \tag{18}$$

The connection between Equations (13) and (18) follows from the law of large numbers

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} F_C\left[\tau(X_i)\right]\left\{Y_i^* - \kappa_C\left[\tau(X_i)\right]\right\} \xrightarrow{a.s.} & \mathbb{E}_{X,Y(1),Y(0)}\Big[F_C\left[\tau(X_i)\right]\left\{Y_i^* - \kappa_C\left[\tau(X_i)\right]\right\}\Big] \\
= & \mathbb{E}_X\Big[F_C\left[\tau(X_i)\right]\left\{\tau_0(X_i) - \kappa_C\left[\tau(X_i)\right]\right\}\Big],
\end{aligned}
$$

where the second line holds because the transformed outcomes, Y_i^*, provide unbiased estimates of the CATEs.

We next discuss two key considerations for implementing our proposed estimator: the surrogate distribution $F_C(\cdot)$ and the parameterization of $\tau(\cdot)$.

4.1 Surrogate Distribution $F_C(\cdot)$

Proposition 2 suggests choosing surrogate distributions $F_C(\cdot)$ that are strictly positive in a neighborhood of $\tau_0(X)$ for all $X \in \mathbb{X}$. Since $\tau_0(X)$ is unknown, we can use distributions with full support. Alternatively, one may employ broad support distributions, that is, distributions whose density is positive over a sufficiently wide domain to cover plausible ranges of $\tau_0(X)$.

We provide analytical expressions for the surrogate objective using normal, logistic, and uniform distributions. These distributions allow an analytical form for the term $\kappa_C(\tau) \equiv \mathbb{E}\left[C | C \leq \tau\right]$. All derivations are provided in Appendix C.

Normal Distribution

Let $F_C(\cdot)$ correspond to a normal distribution with mean c and standard deviation σ, i.e., $C \sim N(c, \sigma^2)$. The normal distribution yields the following estimator

$$
\begin{aligned}
\bar{\tau}(\cdot) &= \underset{\tau:\mathbb{X}\to\mathbb{R}}{\arg\max}\, \frac{1}{n}\sum_{i=1}^{n} \Phi\left[\tau(X_i)\right]\left(Y_i^* - c\right) + \sigma\phi\left[\tau(X_i)\right] \\
\widehat{\tau}(X) &= \bar{\tau}(X)\sigma + c, \text{ for all } X \in \mathbb{X}
\end{aligned}
\tag{19}
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and cdf of the standard normal distribution, respectively. This is the distribution that we use in all our simulations and empirical exercises.

Logistic Distribution

Let $F_C(\cdot)$ correspond to a logistic distribution with mean c and scale σ, i.e., $C \sim \text{Logistic}(c, \sigma)$.

We obtain the following form for the proposed estimator

$$\begin{aligned}\bar{\tau}(\cdot) &= \underset{\tau:\mathbb{X}\to\mathbb{R}}{\arg\max} \frac{1}{n}\sum_{i=1}^{n} G\left[\tau(X_i)\right]\left(Y_i^* - c\right) + \sigma\left(\tau(X_i)\left\{1 - G\left[\tau(X_i)\right]\right\} - \ln G\left[\tau(X_i)\right]\right) \\ \widehat{\tau}(X) &= \bar{\tau}(X)\sigma + c, \text{ for all } X \in \mathbb{X}\end{aligned} \tag{20}$$

where $G(u)$ is the cdf of the standard logistic distribution, that is, $G(u) = \frac{1}{1+e^{-u}}$. One slight advantage of this specification over the normal is that we can express it in closed-form.

Uniform Distribution

Let $F_C(\cdot)$ correspond to the cdf of the uniform distribution between $\underline{c}$ and $\overline{c}$, i.e., $C \sim \mathcal{U}(\underline{c}, \overline{c})$. The proposed estimator becomes

$$\begin{aligned}\widehat{\tau}(\cdot) &= \underset{\tau:\mathbb{X}\to\mathbb{R}}{\arg\max} \frac{1}{n}\sum_{i=1}^{n} 2Y_i^*\tau(X_i) - \left[\tau(X_i)\right]^2 \\ &= \underset{\tau:\mathbb{X}\to\mathbb{R}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \left[Y_i^* - \tau(X_i)\right]^2. \end{aligned} \tag{21}$$

Note that this estimation objective does not depend on the values of $\underline{c}$ and $\overline{c}$.

Discussion

Equation (21) demonstrates a connection between our proposed method and MSE-based approaches to estimate CATEs in Equation (5). When we use the uniform distribution for $F_C(\cdot)$, the two approaches become equivalent. Intuitively, the uniform distribution makes different draws of the stochastic cost C equally likely, and so the estimation errors for CATEs are equally "penalized" in our proposed surrogate objective function. In contrast, the normal and logistic distributions draw C closer to the treatment cost c, which can improve estimation around the decision boundary.

The scale parameter σ in the normal and logistic examples controls how much the proposed objective function deviates from the true profit optimization objective. When σ is small, the modified objective function closely approximates the policy value function from Equation (7). As $\sigma \downarrow 0$, our proposed method becomes closer to a direct policy optimization approach. In contrast, for large σ, the modified objective function can substantially deviate from the profit optimization objective and, as $\sigma \to \infty$, it converges to the MSE-based CATE estimation, as in the uniform example. We treat it as a hyperparameter that needs to be tuned using cross-validation.

4.2 Specification of $\tau(\cdot)$

We parameterize the CATE function $\tau(\cdot)$ with a vector θ and optimize the surrogate objective with respect to θ, that is,

$$\widehat{\theta} = \underset{\theta \in \Theta_n}{\arg\max} \frac{1}{n} \sum_{i=1}^{n} F_C \left[\tau(X_i;\theta) \right] \left\{ Y_i^* - \kappa_C \left[\tau(X_i;\theta) \right] \right\} \tag{22}$$

where Θ_n is the parameter space. The subscript n indicates that it can be a function of the sample size. Notice that our estimator is an M-estimator because we can express the objective function as

$$Q_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} q_i(\theta) \tag{23}$$

where

$$q_i(\theta) \equiv F_C \left[\tau(X_i;\theta) \right] \left\{ Y_i^* - \kappa_C \left[\tau(X_i;\theta) \right] \right\}. \tag{24}$$

With a proper choice of $F_C(\cdot)$, we can obtain a closed-form expression for the derivative $\partial q_i / \partial \theta$, which allows the use of gradient-based optimization algorithms to compute $\widehat{\theta}$. We provide the first- and second-order derivatives for the normal, logistic, and uniform distributions in Appendix D.

Next, we consider two specifications for $\tau(X_i;\theta)$ and briefly discuss their properties.

4.2.1 General Parametric

The simplest approach is to choose a specific parametric functional form, such as linear $\tau(X;\theta) = X^\top \theta$. The linear specification is attractive in empirical applications because it provides an easily interpretable characterization of the CATEs. Under technical conditions, the resulting M-estimator is $\sqrt{n}$-consistent with a limiting normal distribution whose variance is easily estimable. We formalize this in Propositions 3, 4 and 5. The proofs are provided in Appendix E and consist of demonstrating the conditions of Theorems 2.1, 3.1, and 4.1 of Newey and McFadden (1994).

Proposition 3. *Consistency*

Assume that (i) the data are i.i.d.; (ii) Θ is compact; (iii) $|\mathbb{E}[Y_i^*]| \equiv |\mu_*| < \infty$; (iv) $F_C(\cdot)$ is continuous; (v) $\tau(\cdot)$ is continuous in θ; (vi) $\left|\int_{-\infty}^{+\infty} u f_C(u)du\right| \equiv |\mu_C| < \infty$; (vii) and $Q(\theta) \equiv \mathbb{E}\left[Q_n(\theta)\right]$ is uniquely maximized at θ_0. Then $\widehat{\theta} \xrightarrow{p} \theta_0$.

Proposition 3 establishes consistency of our estimator. The basic conditions on the data for it to hold, that the observations are i.i.d. and that Y_i^* has a finite first moment, are plausible in many marketing experiments. Requiring that the distribution $F_C(\cdot)$ is continuous with a finite first moment is trivial because the practitioner chooses this distribution, and the same applies to continuity of $\tau(\cdot)$ in θ. Uniqueness of the maximizer θ_0 is ensured by Proposition 2 when $\tau(\cdot)$ is correctly specified; otherwise, it is maintained as an identifying assumption.

Proposition 4. *Asymptotic Normality*
Assume that the conditions of Proposition 3 hold, so that $\widehat{\theta} \xrightarrow{p} \theta_0$. Define $M(\theta) \equiv \mathbb{E}\left[\frac{\partial q_i(\theta)}{\partial \theta}\frac{\partial q_i(\theta)}{\partial \theta^\top}\right]$ and $B(\theta) \equiv \mathbb{E}\left[\frac{\partial^2 q_i(\theta)}{\partial\theta\partial\theta^\top}\right]$. In addition, assume that (i) $\theta_0 \in interior(\Theta)$; (ii) $F_C(\cdot)$ and $\tau(X_i;\theta)$ are C^2; (iii) $M \equiv M(\theta_0) < \infty$; (iv) $\left|\left|\frac{\partial^2 q_i(\theta)}{\partial\theta\partial\theta^\top}\right|\right| \leq d_B(Y_i^*, X_i)$ with $\mathbb{E}\left[d_B(Y_i^*, X_i)\right] < \infty$; and (v) $B \equiv B(\theta_0)$ is nonsingular. Then it follows that $\sqrt{n}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, B^{-1}MB^{-1}\right)$.

Proposition 4 derives the asymptotic distribution of our estimator, which is normal and has a square root convergence rate. We present it in a general form as a function of $q_i(\cdot)$ because the required conditions might change depending on the choice of $F_C(\cdot)$ and $\tau(\cdot)$. However, they are likely to be weak and boil down to the existence of low-order moments of known functions of the data and sufficient smoothness and boundedness of $q_i(\cdot)$.

Proposition 5. *Consistent Estimator for Covariance Matrix*
Define $\widehat{B} \equiv \frac{1}{n}\sum_{i=1}^n \frac{\partial^2 q_i(\widehat{\theta})}{\partial\theta\partial\theta^\top}$ and $\widehat{M} \equiv \frac{1}{n}\sum_{i=1}^n \frac{\partial q_i(\widehat{\theta})}{\partial\theta}\frac{\partial q_i(\widehat{\theta})}{\partial\theta^\top}$. Assume that the conditions of Propositions 3 and 4 hold. In addition, assume that (i) $\left|\left|\frac{\partial q_i(\theta)}{\partial\theta}\frac{\partial q_i(\theta)}{\partial\theta^\top}\right|\right| \leq d_M(Y_i^*, X_i)$ and (ii) $\mathbb{E}\left[d_M(Y_i^*, X_i)\right] < \infty$. Then $\widehat{B}^{-1}\widehat{M}\widehat{B}^{-1} \xrightarrow{p} B^{-1}MB^{-1}$.

Finally, Proposition 5 presents a consistent estimator for the asymptotic variance from Proposition 4. The same comments with respect to Proposition 4 apply.

The theory of M-estimation for the parametric case is vast and well-established. It is important to note that the conditions we outlined are sufficient for our estimator to possess the usual desired properties, and they can be relaxed as discussed in Newey and McFadden (1994). For example, differentiability of $q_i(\cdot)$ is required only in a neighborhood of θ_0

and not throughout the entire space Θ. We opted to present the results in a simpler way because they are likely to hold in our setting.

4.2.2 Sieves

We can use sieve estimators to approximate CATEs using a flexible functional form. Sieve M-estimators use progressively complex models to estimate an unknown high-dimensional function as the sample size increases, so that in the limit the resulting estimator recovers the true underlying function. We direct the reader to Chen (2007) for a comprehensive review of sieve estimators.

A particularly attractive approach to specify $\tau(X_i;\theta)$ is via neural networks (NNs). Although they offer less interpretability, NNs can provide more flexibility and better fit in estimating CATEs. They can be trained using standard packages with the custom loss function defined in Equation (22). The backpropagation approach to model training requires specifying the derivative for the final layer, which depends on the choice of the distribution $F_C(\cdot)$. We provide them for the uniform, normal, and logistic distributions in Appendix D.

In practical applications, incorporating regularization is important for both out-of-sample performance and numerical stability. For the linear specification, we apply ℓ_1 regularization to the model coefficients. For the neural network specification, we use standard regularization techniques, including weight decay, dropout, gradient clipping, and early stopping. Batch normalization is also employed to stabilize optimization and accelerate convergence.

5 Illustrative Applications

We use synthetic data to demonstrate the theoretical tradeoffs and performance of the proposed estimator. We consider two data generating processes (DGPs). The first DGP illustrates the trade-off between MSE-focused estimation and performance optimization using CATEs that imply a single-threshold decision rule. The second DGP considers a different setup with a more-complex decision rule, which allows us to contrast the proposed estimator to the advanced targeting methods.

5.1 Simulation: Simple CATEs

We begin with the following specification for the CATEs

$$\begin{aligned} Y_i &= 1 + X_i + W_i \tau_0(X_i) + \epsilon_i \\ \tau_0(X_i) &= -X_i^2 + 2X_i + 1 \end{aligned} \tag{25}$$

where $X_i \sim \mathcal{U}(-1, 2)$, $W_i \sim$ Bernoulli(0.5), and $\epsilon_i \sim N(0, 0.1)$. We assume that the treatment cost is equal to one, $c = 1$, which implies that the optimal decision rule is to mail customers if and only if $X \geq 0$. We depict the true CATEs and the decision boundary in Figure 2.

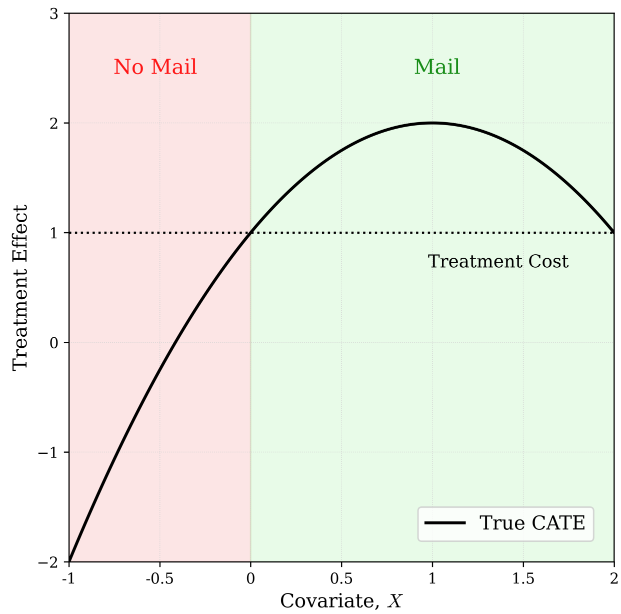

Figure 2: Simple CATEs and the Corresponding Targeting Policy

How well does our proposed method perform in recovering the CATEs and the optimal decision rule? Although there are theoretical guarantees that the estimator, when correctly specified, will be consistent, whether it performs well in finite samples is an empirical question. To evaluate this, we simulate samples of 10,000 observations and estimate the proposed method with $\tau(X_i; \theta) = \theta_0 + \theta_1 X_i + \theta_2 X_i^2$ and using $C \sim N(c, 1)$. In Figure 3, we compare the estimated CATEs to the true CATEs from the DGP. The dashed red line shows the mean of the estimated CATEs across 1,000 iterations, and the shaded area represents the 95% confidence interval. The results confirm that with a correct specification, our proposed estimator closely approximates the true CATEs. Estimation accuracy is highest near the decision boundary ($X = 0$), whereas variance increases for observations farther from the boundary.

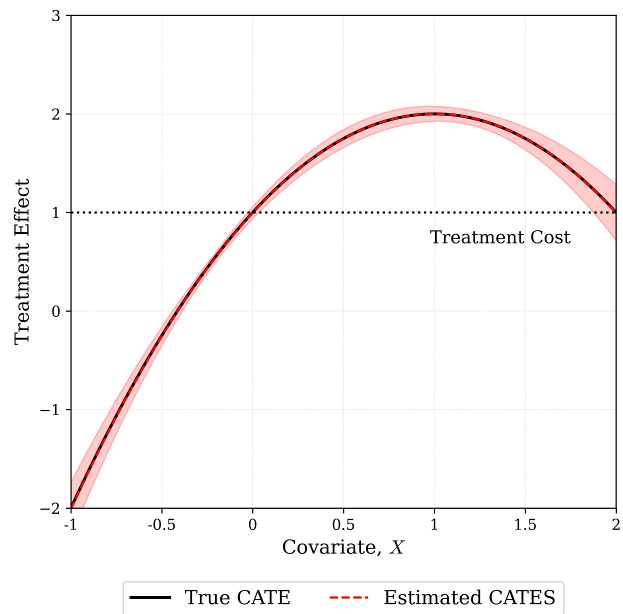

Figure 3: True and Estimated CATEs

A different matter is the performance of our estimator when it is misspecified. We consider this scenario by applying our proposed method with a linear specification $\tau(X_i;\theta) = \theta_0 + \theta_1 X_i$ to the same (quadratic) DGP. Figure 4 shows how the proposed estimator performs with different values of the parameter σ for a single draw of the training data. It plots the true CATEs, the linear approximation obtained from our estimator, and highlights the region where the implied targeting decisions are different from optimal.

Figure 4 illustrates the trade-off between profit optimization and prediction accuracy. It demonstrates how this trade-off plays out in practice. With a large σ, our method performs similarly to estimating OLS with transformed outcomes, fitting the true CATEs uniformly over the full range of X. As we decrease σ, the proposed method places more weight on profit optimization. The estimated CATEs are closer to the true CATEs near the decision boundary, but errors are larger for the extreme values of X. With very small σ, the objective function approaches a pure profit maximization objective. The error region shrinks, and we observe that the estimated line is almost tangential to the true CATE at $X = 0$.

We further explore this trade-off by contrasting the accuracy of CATE estimation (MSE) and a value of the implied targeting policy (Profit). In expectation, OLS yields the smallest MSE among the linear models, so the question is how much prediction accuracy our

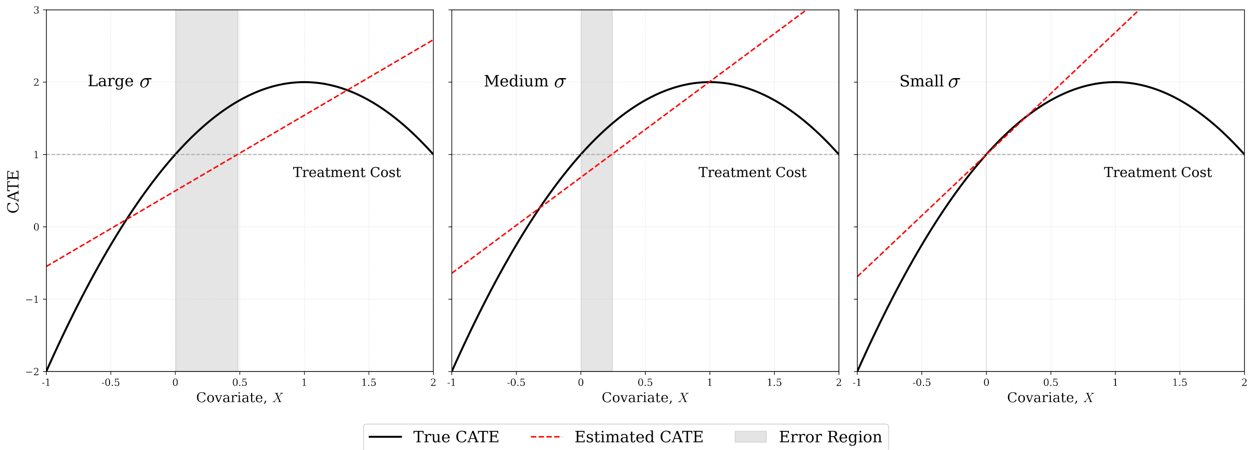

Figure 4: Estimated CATEs with Linear Specification and $F_C(\cdot) = \mathcal{N}(c, \sigma)$

method would sacrifice to obtain higher incremental profits. For ease of visualization, Figure 5 plots the *negative* MSE against the incremental profits captured by our proposed method with different values of σ in $C \sim N(c, \sigma^2)$. We demonstrate the model performance for a single draw of the training data with 10,000 observations.

Figure 5 shows that linear models can exceed the incremental profit obtained by OLS at the expense of MSE. The tradeoff is controlled by the hyperparameter σ. The optimal choice of this parameter depends on one's preferences in a particular application. The curve shown in Figure 5 can be seen as a "possibility frontier" that shows the limit of what can be obtained based on the data under the linear specification of $\tau(X)$. Notice that we can further improve the model performance on both incremental profits and estimation accuracy when using the correct specification: with quadratic $\tau(X)$, our method approaches the highest performance attainable under this DGP (Oracle).

5.2 Simulation: Complex Decision Boundary

We now consider a different DGP that features nonlinear CATEs and a more complex optimal decision rule:

$$\begin{aligned} Y_i &= W_i \tau_0(X_i) + \epsilon_i \\ \tau_0(X_i) &= \omega^\top X_i \sin(2.3 \cdot \omega^\top X_i) + 1.3 \end{aligned} \tag{26}$$

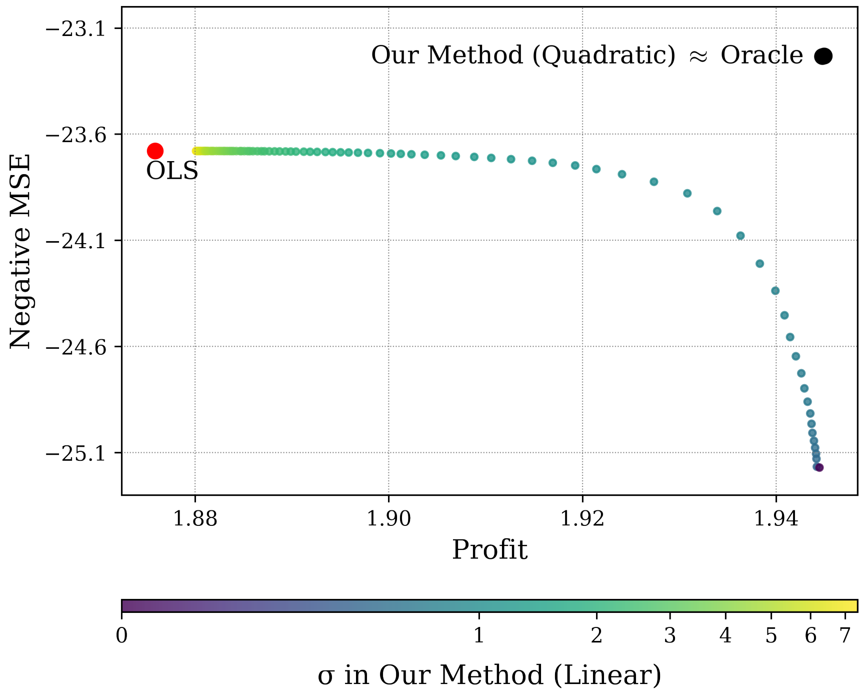

Figure 5: Trade-off Between MSE and Incremental Profits

where

$$
\begin{aligned}
X_i &= (X_{i1}, X_{i2}, \ldots, X_{i,10})^\top \\
X_{ij} &\overset{\text{i.i.d.}}{\sim} \text{Uniform}(-1, 2), \quad j \in 1, \ldots, 10 \\
\omega &= \frac{1}{\sqrt{10}}(1, 1, \ldots, 1)^\top \in \mathbb{R}^{10}
\end{aligned}
\tag{27}
$$

We further assume $W_i \sim \text{Bernoulli}(0.5)$, $\epsilon_i \sim N(0, 0.1)$, and $c = 1$.

Figure 6 illustrates this DGP. Compared to Section 5.1, this specification introduces nonlinearities in the covariates and features no additive separability. The optimal decision rule now includes multiple cutoffs in the covariate index $z_i = \omega^\top X_i$. Consequently, nonlinear models are expected to outperform linear baselines in both estimating the targeting policy and recovering the CATEs.

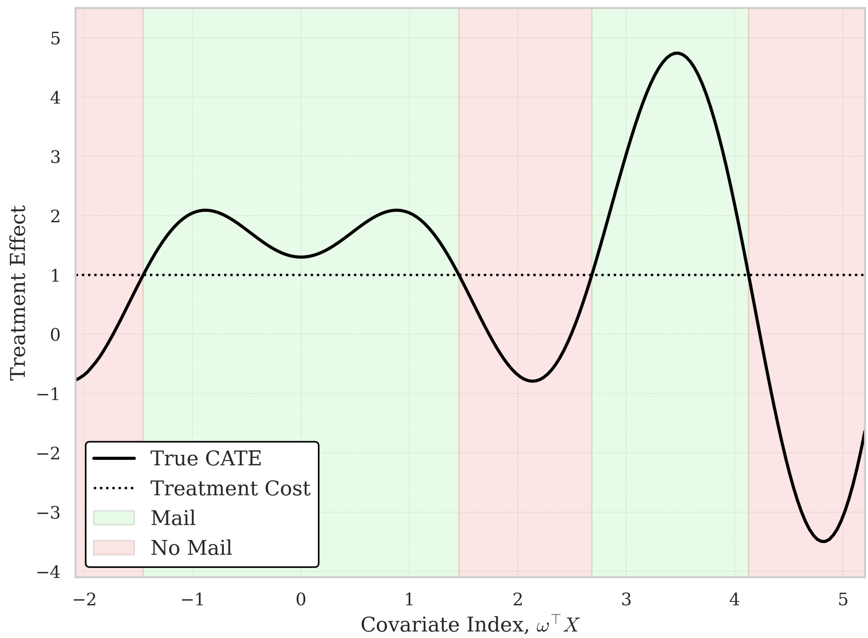

Figure 6: CATE and Complex Optimal Targeting Policy

We compare the performance of the proposed method to several baselines: OLS, direct policy optimization using a deep neural network (PolicyDNN; Zhang, 2025), gradient-boosting (XGBoost), and causal forest (Wager and Athey, 2018). We estimate OLS and XGBoost using the transformed outcomes approach, as described in Equation (5). We calibrate each method using 10,000 observations and then evaluate the estimated treatment effects and targeting policies using the true DGP. For our method, we specify the surrogate distribution as normal centered at the treatment cost, $\mathcal{N}(c, \sigma^2)$, and use the cross-validation approach to choose two values of σ: one that minimizes MSE and one that maximizes profit. The cross-validation grid includes $\sigma = \infty$, which is equivalent to fitting transformed outcomes with a neural network. We average the model performance across 100 draws of the training data.

We report the results in Table 2. The table includes three outcome measures: (i) incremental *Profit* relative to a blanket no-mail policy, (ii) *MSE* between the estimated and true CATEs from the DGP, and (iii) the *Qini* coefficient, which measures the quality of customer ranking implied by the estimated CATEs (Yadlowsky et al., 2025). Best methods yield large *Profit* and *Qini*, and small MSE.

All targeting methods outperform the uniform baselines. As expected, the linear specification of our method performs similarly to OLS when σ is selected to minimize MSE. How-

Table 2: Method Comparison with Complex Simulation

	Profit		**MSE**		**Qini**	
Model	Mean	SD	Mean	SD	Mean	SD
Uniform Baselines						
Mail	-0.004	–	–	–	0.000	–
No Mail	0.000	–	–	–	0.000	–
Linear Models						
OLS	0.268	0.026	1.496	0.003	0.096	0.004
Our method (linear, MSE)	0.274	0.023	1.495	0.003	0.097	0.004
Our method (linear, Profit)	0.312	0.040	2.354	0.287	0.102	0.012
Tree-Based Methods						
Causal Forest	0.280	0.022	1.394	0.004	0.187	0.008
XGBoost	0.271	0.005	1.175	0.013	0.185	0.003
Neural Networks						
PolicyDNN	0.512	0.001	–	–	0.278	0.015
Our method (NNet; MSE)	0.509	0.004	0.088	0.033	0.336	0.004
Our method (NNet; Profit)	0.511	0.001	0.392	0.836	0.328	0.017
CATEs from DGP						
Oracle	0.515	–	0.000	–	0.344	–

ever, our method allows to shift emphasis from global CATE accuracy toward decision-relevant precision near the boundary. When we choose the profit-maximizing σ in cross-validation, the linear version of our method achieves substantially higher profit than OLS, and even outperforms more flexible tree-based approaches, including Causal Forest and XGBoost.

Although tree-based methods achieve lower MSE than any of the linear specifications, this precision does not necessarily translate into superior profit. This result underscores the central insight of our work: improving policy value requires accurate CATE estimates specifically near the decision boundary, not uniformly across the entire covariate space.[4]

Neural network methods yield targeting policies with near-optimal profit. Recall that PolicyDNN focuses solely on profit maximization and does not estimate CATEs. Table 2 further demonstrates that PolicyDNN scores are not reliable proxies for CATEs; they produce notably worse customer ranking (Qini) than both our method and the Oracle benchmark. In contrast, our method modifies the direct profit optimization problem by introducing a principled surrogate objective that enables CATE estimation along with the targeting rule. This surrogate introduces a larger discrepancy from the stepwise profit function than the softmax transformation used in PolicyDNN, but the difference does not substantially reduce profits in this setting.

Overall, the results in Table 2 highlight the flexibility of our method. By selecting an appropriate value of σ, practitioners can navigate the trade-off between global CATE accuracy and profit maximization. In settings that involve high-dimensional nonlinear treatment effects with interactions terms, this flexibility yields notable gains in policy value while still providing interpretable policy-aligned CATE estimates that rationalize the resulting targeting decisions.

6 Conclusion

In this paper, we proposed a methodology to estimate CATEs that simultaneously optimizes expected payoffs. By reframing treatment effect estimation as a profit maximization problem with a stochastic threshold, our framework reconciles the two dominant paradigms in targeting policy design: direct policy optimization and plug-in estimation. Our findings underscore that global estimation accuracy is insufficient for maximizing

[4] In our simulation, tree-based methods can almost perfectly approximate CATEs and yield near-optimal decision rules with one-dimensional covariates, $\dim(X_i) = 1$. However, their performance deteriorates when the CATE function involves complex interactions between covariates in multiple dimensions.

policy performance, yet targeting decisions made without CATEs lack the interpretability necessary for strategic oversight. We derive the statistical properties of the proposed estimation approach and illustrate the theoretical tradeoffs using synthetic data. Our approach has broad relevance across domains, ranging from personalized medicine to operations research, where firms and policymakers must balance the efficiency of targeting policies with the need for transparent, explainable decision rules.

Limitations and Future Research

Our work opens several avenues for future research to address current limitations and extend the framework's applicability. First, while our current approach focuses on binary conditions, many practical applications involve complex decision spaces. Future research can extend the proposed surrogate objective to accommodate multi-valued treatment arms and continuous treatment effects, such as personalized pricing or incrementality-based advertising. This would broaden the applicability of the method beyond the current settings that are tailored to experimental data with binary marketing actions.

Second, it would be valuable to characterize the discrepancy between the proposed surrogate objective and standard profit optimization in finite samples. We have established that our surrogate objective function is Fisher consistent with respect to the original profit function, ensuring asymptotic validity. However, in finite samples, the surrogate objective may deviate from the true profit maximization goal. Future research could derive theoretical bounds to formally characterize this discrepancy.

Third, the reliance on experimental data points to important extensions regarding observational data and the unconfoundedness assumption. When using observational data, propensity scores must be estimated, and the impact of this first estimation step on the resulting CATE estimator and decision-making warrants further investigation. Additionally, future work could investigate how to adapt the framework when the unconfoundedness assumption is not defensible, perhaps by incorporating instrumental variables.

Finally, our method could be integrated with adaptive experimental designs to enhance learning efficiency. For instance, firms could adopt sampling procedures similar to Chen et al. (2025) to adaptively increase sample sizes based on customer characteristics. This integration would allow for maximizing learning about CATEs in regions critical to the optimal policy, potentially requiring a better characterization of variance at the individual customer level or an extension of the proposed approach into a Bayesian framework.

References

Ascarza, E. and Israeli, A. (2022). Eliminating unintended bias in personalized policies using bias-eliminating adapted trees (BEAT). *Proceedings of the National Academy of Sciences*, 119(11):e2115293119.

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. and Wager, S. (2021). Policy learning with observational data. *Econometrica*, 89(1):133–161.

Blattberg, R. C. and George, E. I. (1992). Estimation under profit-driven loss functions. *Journal of Business Economics & Statistics*, 10(4):437–444.

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics, Vol. 6B*, pages 5549–5632. North-Holland.

Chen, Y.-W., Ascarza, E., and Netzer, O. (2025). Policy-aware experimentation: Strategic sampling for optimized targeting policies. SSRN Working Paper 5044549.

Dahabreh, I. J., Hayward, R., and Kent, D. M. (2016). Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. *International Journal of Epidemiology*, 45(6):2184–2193.

Dehejia, R. H. (2005). Program evaluation as a decision problem. *Journal of Econometrics*, 125(1–2):141–173.

Dubé, J.-P. and Misra, S. (2023). Personalized pricing and consumer welfare. *Journal of Political Economy*, 131(1):131–189.

Elmachtoub, A. N. and Grigas, P. (2022). Smart "predict, then optimize". *Management Science*, 68(1):9–26.

Fernández-Loría, C. and Provost, F. (2022). Causal decision making and causal effects estimation are not the same... and why it matters. *INFORMS Journal on Data Science*, 1(1):4–16.

Hirano, K. and Porter, J. R. (2020). Asymptotic analysis of statistical decision rules in econometrics. In Durlauf, S. N., Hansen, L. P., Heckman, J. J., and Matzkin, R. L., editors, *Handbook of Econometrics, Vol. 7A*, pages 283–354. North-Holland.

Hitsch, G. J., Misra, S., and Zhang, W. W. (2024). Heterogeneous treatment effects and optimal targeting policy evaluation. *Quantitative Marketing and Economics*, 22(2):115–168.

Hu, A. (2023). Heterogeneous treatment effects analysis for social scientists: A review. *Social Science Research*, 109.

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

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1):4–29.

Jacob, D. (2021). CATE meets ML: Conditional average treatment effect and machine learning. *Digital Finance*, 3(2):99–148.

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

Künzel, S., Sekhon, J. S., Bickel, P. J., and Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10):4156–4165.

Lemmens, A. and Gupta, S. (2020). Managing churn to maximize profits. *Marketing Science*, 39(5):956–973.

Lin, Y. (2004). A note on margin-based loss functions in classification. *Statistics & Probability Letters*, 68(1):73–82.

Lu, H., Simester, D., and Zhu, Y. (2025). Optimizing scalable targeted marketing policies with constraints. *Marketing Science*, 44(5):1082–1103.

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

Mbakop, E. and Tabord-Meehan, M. (2021). Model selection for treatment choice: Penalized welfare maximization. *Econometrica*, 89(2):825–848.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics, Vol. 4*, pages 2111–2245. North-Holland.

Nie, X. and Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319.

Rafieian, O. and Yoganarasimhan, H. (2021). Targeting and privacy in mobile advertising. *Marketing Science*, 40(2):193–218.

Simester, D., Timoshenko, A., and Zoumpoulis, S. I. (2020a). Efficiently evaluating targeting policies: Improving on champion vs. challenger experiments. *Management Science*, 66(8):3412–3424.

Simester, D., Timoshenko, A., and Zoumpoulis, S. I. (2020b). Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Science*, 66(6):2495–2522.

Swaminathan, A. and Joachims, T. (2015). Batch learning from logged bandit feedback through counterfactual risk minimization. *Journal of Machine Learning Research*, 16(1):1731–1755.

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.

Waisman, C., Nair, H. S., and Carrion, C. (2025). Online causal inference for advertising in real-time bidding auctions. *Marketing Science*, 44(1):176–195.

Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., and Wager, S. (2025). Evaluating treatment prioritization rules via rank-weighted average treatment effects. *Journal of the American Statistical Association*, 120(549):38–51.

Zhang, W. W. (2025). Optimal comprehensible targeting. Working Paper, University of Pennsylvania.

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

Zhou, H., Li, S., Jiang, G., Zheng, J., and Wang, D. (2023). Direct heterogeneous causal learning for resource allocation problems in marketing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 5446–5454.

Appendix

A Derivations from Section 3.1

We now demonstrate how Equation (12) is equivalent to Equation (13).

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}\left\{\tau \geq C\right\}(\tau_0 - C)\right] &= \int_{-\infty}^{\tau} (\tau_0 - u)\, dF_C(u)du \\
&= F_C(\tau)\left(\tau_0 - \int_{-\infty}^{\tau} \frac{u}{F_C(\tau)} dF_C(u)du\right) \\
&= F_C(\tau)\left(\tau_0 - \mathbb{E}\left[C|C \leq \tau\right]\right) \\
&\equiv F_C(\tau)\left[\tau_0 - \kappa_C(\tau)\right].
\end{aligned}$$

B Additional Illustration of the Surrogate Objective Function

In Figure B.1, we illustrate the surrogate objective function for $C \sim N(c, \sigma^2)$ and varying values of σ. The simulation assumes $c = 1$ and $\tau_0 = 2$. The surrogate objective function deviates from the true profit maximization objective more for larger σ's. As variance σ decreases, the surrogate objective approaches the stepwise function.

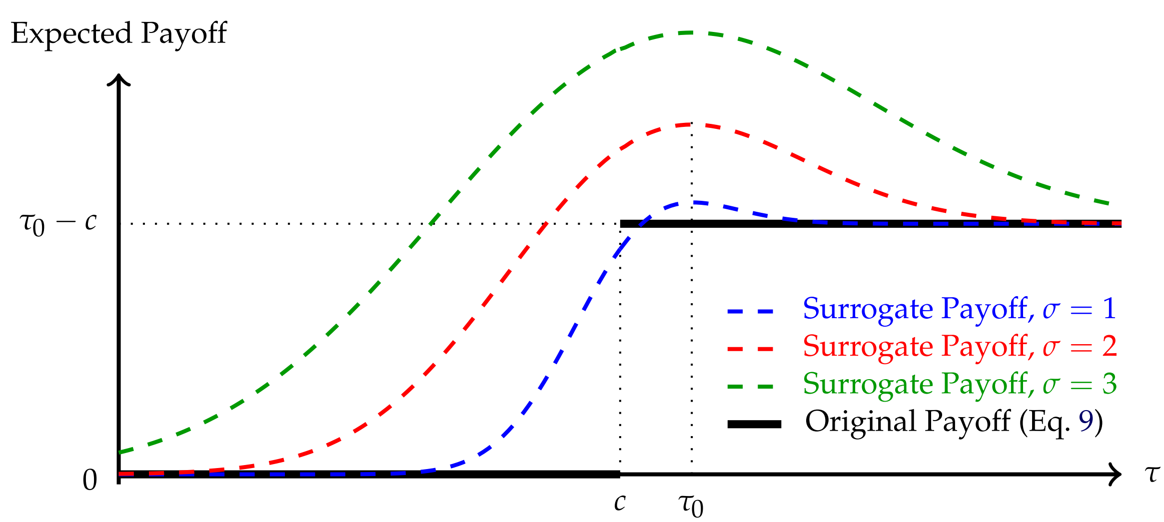

Figure B.1: Illustration of the Surrogate Objective Function for Different Values of σ.

C Derivations from Section 4

C.1 Normal Distribution

We first rewrite the expression of interest

$$\int_{-\infty}^{\tau(X_i)} [Y_i^* - u] f(u) du = \int_{-\infty}^{\tau(X_i)} [Y_i^* - c + c - u] f(u) du$$
$$= (Y_i^* - c) \int_{-\infty}^{\tau(X_i)} f(u) du + \sigma \int_{-\infty}^{\tau(X_i)} \frac{u - c}{\sigma} f(u) du. \tag{C.1}$$

Define $\tilde{C} = \frac{C-c}{\sigma}$, so that $\tilde{C}$ follows a standard normal distribution. The first integral in (C.1) is

$$\int_{-\infty}^{\tau(X_i)} f(u) du = \Pr(C < \tau(X_i)) = \Pr\left(\tilde{C} < \frac{\tau(X_i) - c}{\sigma}\right) = \Phi\left(\frac{\tau(X_i) - c}{\sigma}\right). \tag{C.2}$$

Now consider the second integral

$$\int_{-\infty}^{\tau(X_i)} \frac{u - c}{\sigma} f(u) du = \sigma \int_{-\infty}^{\frac{\tau(X_i) - c}{\sigma}} \tilde{c} f(\sigma \tilde{c} + c) d\tilde{c} = \int_{-\infty}^{\frac{\tau(X_i) - c}{\sigma}} \tilde{c} \phi(\tilde{c}) d\tilde{c} = \phi\left(\frac{\tau(X_i) - c}{\sigma}\right). \tag{C.3}$$

Plugging (C.2) and (C.3) back into (C.1) yields Equation (19).

C.2 Logistic Distribution

Now assume that $f(\cdot)$ is the density of a logistic distribution with location parameter c and scale parameter σ. We now have

$$
\begin{aligned}
\int_{-\infty}^{\tau(X_i)} (Y_i^* - c) f(u) du &= Y_i^* F\left[\tau(X_i\right] - \left[uF(u)\Big|_{-\infty}^{\tau(X_i)} - \int_{-\infty}^{\tau(X_i)} F(u) du \right] \\
&= Y_i^* F\left[\tau(X_i)\right] - \tau(X_i) F\left[\tau(X_i)\right] + \int_{-\infty}^{\tau(X_i)} F(u) du \\
&= [Y_i^* - \tau(X_i)] F\left[\tau(X_i)\right] + \sigma \int_{-\infty}^{\frac{\tau(X_i)-c}{\sigma}} F(\sigma \tilde{c} + c) d\tilde{c} \\
&= \frac{Y_i^* - \tau(X_i)}{1 + \exp\left\{ -\frac{\tau(X_i)-c}{\sigma} \right\}} + \sigma \int_{-\infty}^{\frac{\tau(X_i)-c}{\sigma}} \frac{1}{1+e^{-\tilde{c}}} d\tilde{c} \\
&= \frac{Y_i^* - \tau(X_i)}{1 + \exp\left\{ -\frac{\tau(X_i)-c}{\sigma} \right\}} + \sigma \log\left(1 + \exp\left\{ \frac{\tau(X_i) - c}{\sigma} \right\} \right).
\end{aligned}
$$

C.3 Uniform Distribution

Assume that $f(\cdot)$ corresponds to the density of a uniform distribution between $\underline{c}$ and $\bar{c}$. Then we have:

$$
\begin{aligned}
\int_{-\infty}^{\tau(X_i)} (Y_i^* - u) f(u) du &= \int_{\underline{c}}^{\tau(X_i)} [Y_i^* - u] \frac{1}{\bar{c} - \underline{c}} du \\
&= \frac{Y_i^*}{\bar{c} - \underline{c}} \int_{\underline{c}}^{\tau(X_i)} du - \frac{1}{\bar{c} - \underline{c}} \int_{\underline{c}}^{\tau(X_i)} u du \\
&= \left(\frac{\tau(X_i) - \underline{c}}{\bar{c} - \underline{c}} \right) Y_i^* - \frac{\tau(X_i)^2 - \underline{c}^2}{2(\bar{c} - \underline{c})}.
\end{aligned}
$$

To see the connection between this expression and the MSE, notice that

$$
\begin{aligned}
\arg\max_{\tau:\mathbb{X}\to\mathbb{R}} \left[\left(\frac{\tau(X_i) - \underline{c}}{\bar{c} - \underline{c}} \right) Y_i^* - \frac{\tau(X_i)^2 - \underline{c}^2}{2(\bar{c} - \underline{c})} \right] &= \arg\max_{\tau:\mathbb{X}\to\mathbb{R}} \left[Y_i^* \tau(X_i) - \frac{1}{2} \tau(X_i)^2 \right] \\
&= \arg\min_{\tau:\mathbb{X}\to\mathbb{R}} \left[-2\tau(X_i) Y_i^* + \tau(X_i)^2 \right] \\
&= \arg\min_{\tau:\mathbb{X}\to\mathbb{R}} \left[Y_i^* - \tau(X_i) \right]^2.
\end{aligned}
$$

D Standard Errors from Central Limit Theorem for Linear CATE

Recall that

$$\widehat{\theta} = \underset{\theta\in\Theta}{\operatorname{argmax}} \frac{1}{n}\sum_{i=1}^{n} F_C\left[\tau\left(X_i;\theta\right)\right]\left\{Y_i^* - \kappa_C\left[\tau\left(X_i;\theta\right)\right]\right\} = \underset{\theta\in\Theta}{\operatorname{argmax}} \frac{1}{n}\sum_{i=1}^{n} q_i(\theta),$$

where $\kappa_C(\tau) = \mathbb{E}\left[C \mid C \leq \tau\right]$ for a distribution $C \sim F_C(\cdot)$. To simplify the exposition, we will denote $\tau_i := \frac{X_i^T\theta - c}{\sigma}$.

D.1 Normal Distribution

From Equation (19), under a normal distribution we now have

$$q_i(\theta) = \Phi\left(\tau_i\right)\left(Y_i^* - c\right) + \sigma\phi\left(\tau_i\right),$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ correspond to the standard normal cdf and pdf, respectively. The relevant derivatives of q are

$$\frac{\partial q_i(\theta)}{\partial\theta} = \frac{\phi(\tau_i)}{\sigma}(Y_i^* - c - \sigma\tau_i)X_i, \quad \frac{\partial^2 q}{\partial\theta\partial\theta^\top} = -\frac{\phi(\tau_i)}{\sigma^2}\left[\tau_i(Y_i^* - c - \sigma\tau_i) + \sigma\right] X_i X_i^\top.$$

D.2 Logistic Distribution

Finally, from Equation (20) a logistic distribution yields

$$q_i(\theta) = G(\tau_i)\left(Y_i^* - c\right) + \sigma\left\{\tau_i\left[1 - G(\tau_i)\right] - \ln G(\tau_i)\right\},$$

where $G(u) = 1/(1+e^{-u})$ and $g(u) = G'(u) = G(x)\left[1 - G(x)\right]$ correspond to the cdf and the pdf of the standard logistic distribution, respectively. The relevant derivatives are

$$\frac{\partial q_i(\theta)}{\partial\theta} = \frac{g(\tau_i)}{\sigma}\left[Y_i^* - c - \sigma\tau_i\right] X_i, \quad \frac{\partial^2 q_i(\theta)}{\partial\theta\partial\theta^\top} = -\frac{g(\tau_i)}{\sigma^2}\left\{\left[2G(\tau_i) - 1\right]\left(Y_i^* - c - \sigma\tau_i\right) + \sigma\right\} X_i X_i^\top.$$

D.3 Uniform Distribution

From Equation (21), under a uniform distribution $q_i(\theta)$ collapses to the mean squared loss, so that

$$q_i(\theta) = -\left[Y_i^* - X_i^T\theta\right]^2.$$

Taking derivatives, we get

$$\frac{\partial q_i(\theta)}{\partial\theta} = 2\left(Y_i^* - X_i^T\theta\right)X_i, \quad \frac{\partial^2 q_i(\theta)}{\partial\theta\partial\theta^\top} = -2X_iX_i^\top.$$

E Proofs of Propositions

E.1 Proof of Proposition 1

Proof. Consider Equation (12). For any $\tau < \tau_0$, the function $(\tau_0 - u)$ is positive and continuous at $u = \tau$, so a small increase in τ increases the optimization objective. Similarly, for any value $\tau > \tau_0$, the optimization objective increases with a small decrease in τ.[5] Because τ_0 also maximizes the original objective function from Equation (9), the proposed surrogate objective function is Fisher consistent with respect to the original one. □

E.2 Proof of Proposition 2

Proof. Because we do not impose any restrictions on $\tau(\cdot)$ when optimizing (16), solving this optimization problem corresponds to maximizing $\tilde{Q}(\tau)$ independently for all $X \in \mathbb{X}$. Proposition 1 then implies that $\tau_0(X)$ maximizes $\tilde{Q}(\tau)$.

The same logic applies to $Q(\tau)$, so its maximizer is $[c, +\infty)$ for all X such that $\tau_0(X) \geq c$ and $(-\infty, c)$ otherwise. As a result, $\tau_0(X)$ also maximizes $Q(\tau)$, and so $\tilde{Q}(\tau)$ is Fisher consistent with respect to $Q(\tau)$. □

E.3 Proof of Proposition 3

Proof. Our proof consists of demonstrating that the conditions of Theorem 2.1 of Newey and McFadden (1994) are met. First, note that continuity of $F_C(\cdot)$ and of $\tau(\cdot)$ in θ imply that $Q(\cdot)$ is also continuous.

[5] The proposed modified optimization problem is inspired by the literature on advertising auctions. Online advertisers often make decisions based on the distribution of competing bids, and for a given bid, the advertiser's outcomes are stochastic rather than deterministic. Waisman et al. (2025) leverage this insight to infer the effectiveness of online advertising. Conceptually, our approach requires a targeting policy to "submit bids" by choosing τ, and customers are only treated if this "bid" exceeds the (stochastic) treatment cost C, which plays the role of the highest competing bid. Similar to Waisman et al. (2025), the optimal bid coincides with the true ATE, τ_0.

Second, it follows that

$$\begin{aligned}
|q_i(\theta)| &= \left| F_C\left[\tau(X_i;\theta)\right] Y_i^* - \int_{-\infty}^{\tau(X_i;\theta)} u f_C(u) du \right| \\
&\leq F_C\left[\tau(X_i;\theta)\right] |Y_i^*| + \left| \int_{-\infty}^{+\infty} u f_C(u) du \right| \\
&\leq |Y_i^*| + |\mu_C| \\
&\equiv d_q(Y_i^*),
\end{aligned}$$

where the first inequality follows from $f_C(\cdot) \geq 0$ and from the triangle inequality, and the second inequality holds because $F_C(\cdot) \leq 1$.

Because we have $\mathbb{E}\left[d_q(Y_i^*)\right] = |\mu^*| + |\mu_C| < \infty$, the conditions of Lemma 2.4 of Newey and McFadden (1994) are met. Consequently, it follows that $Q(\theta)$ is continuous and $\sup_{\theta \in \Theta} |Q_n(\theta) - Q(\theta)| \xrightarrow{p} 0$. Therefore, the conditions of Theorem 2.1 of Newey and McFadden (1994) are also met, and $\widehat{\theta} \xrightarrow{p} \theta_0$. □

E.4 Proof of Proposition 4

Proof. As with Proposition 3, our proof consists of showing that the conditions for Theorem 3.1 of Newey and McFadden (1994) hold. First, note that the conditions from Proposition 3 plus (i)–(iii) imply that the requirements for the Central Limit Theorem are satisfied, so we obtain $\sqrt{n}\frac{\partial Q_n(\theta_0)}{\partial \theta} \xrightarrow{d} N(0, M)$.

Next, (iv) implies that Lemma 2.4 of Newey and McFadden (1994) applies again, so that $B(\theta)$ is continuous and $\sup_{\theta \in \Theta} \left|\left| \frac{1}{n}\sum_{i=1}^n \frac{\partial^2 q_i(\theta)}{\partial \theta \partial \theta^\top} - B(\theta) \right|\right| \xrightarrow{p} 0$. Therefore, the conditions of Theorem 3.1 of Newey and McFadden (1994) are also satisfied, and so it follows that $\sqrt{n}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, B^{-1}MB^{-1}\right)$. □

E.5 Proof of Proposition 5

Proof. The conditions from Propositions 3 and 4 already yield $\widehat{B}^{-1} \xrightarrow{p} B^{-1}$ due to Lemma 2.4 of Newey and McFadden (1994) and the continuous mapping theorem.

Conditions (i) and (ii) further imply that we can apply Lemma 2.4 of Newey and McFadden (1994) again, so that $M(\theta)$ is continuous and $\sup_{\theta \in \Theta} \left|\left| \frac{1}{n}\sum_{i=1}^n \frac{\partial q_i(\theta)}{\partial \theta}\frac{\partial q_i(\theta)}{\partial \theta^\top} - M(\theta) \right|\right| \xrightarrow{p} 0$, which ensures that $\widehat{M} \xrightarrow{p} M$. Applying the continuous mapping theorem and Slutsky's theorem establishes the result. $\square$