Policy-Aware Experimentation: Strategic Sampling for Optimized Targeting Policies

Yi-Wen Chen
Columbia Business School, YChen26@gsb.columbia.edu

Eva Ascarza
Harvard Business School, eascarza@hbs.edu

Oded Netzer
Columbia Business School, onetzer@gsb.columbia.edu

With unprecedented access to consumer information, firms are increasingly interested in designing highly effective data-driven targeting policies based on detailed consumer data. The current standard for implementing such policies involves the "test-then-learn" approach, where randomized experiments are used to estimate the differential impact of marketing interventions on various customers. However, this method fails to incorporate the firm's ultimate business objectives, leading to inefficient experimentation and suboptimal targeting strategies. To overcome this limitation, we propose a sequential experimental design integrated with a novel sampling criterion — expected profit loss — which aligns theoretically with the firm's profit-maximizing objective. Additionally, we introduce a novel expected profit loss estimation method leveraging the power of Bayesian inference for uncertainty quantification based on Causal Forest. Through extensive simulation studies and two empirical applications, we demonstrate the superiority of our approach in improving targeting performance. Furthermore, we emphasize the effectiveness of our approach even when simplified into a two-stage design, enabling firms to shorten the experimentation period and streamline the process. Our research underscores the importance of aligning experimental design with business objectives and offers an efficient solution for firms seeking to enhance their targeting strategies.

Key words: Policy learning, marketing interventions, targeted policies, experimentation, active learning, heterogeneous treatment effect

History: This version: Dec 03, 2024

---

1. Introduction

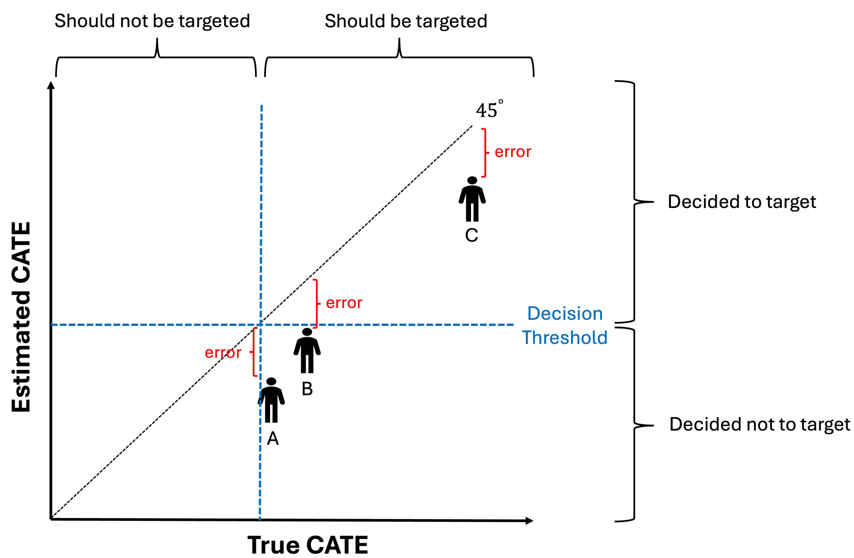
In the digital era, firms have unprecedented access to consumer information, enabling the design of highly effective data-driven targeting policies based on detailed consumer data. These policies rely on identifying which customers are most likely to respond favorably to marketing interventions, allowing firms to tailor their strategies to maximize business outcomes (e.g., Ascarza 2018, Simester et al. 2020, Yang et al. 2023). However, the effectiveness of these strategies depends on how well firms can predict customer responsiveness, which, in turn, is determined by how they sample customers during the experimentation phase. Most experimentation sampling strategies used today are based on statistical principles and implementation costs, focusing on customer representativeness (within a budget). Such strategies tend to ignore the firm's business objectives when they decide who to sample. This misalignment often results in suboptimal targeting policies that fail to maximize profitability (Fernández-Loría and Provost 2022).

Consider the current common practice in marketing, the "test-then-learn" approach. This method involves conducting randomized experiments to estimate how customers respond to interventions, with the goal of developing targeting policies based on these predictions (e.g., Lemmens and Gupta 2020, Ellickson et al. 2022, Huang and Ascarza 2024). Typically, this approach follows three steps. First, during the *experimentation* phase, firms sample a subset of customers from their customer base according to a predetermined experimental size and randomly assign them to treatment or control groups. Next, in the *estimation* phase, firms build models to estimate the average treatment effect (ATE) and the conditional average treatment effect (CATE), representing the treatment effect conditioned on customer characteristics. Finally, in the *policy implementation* phase, targeting decisions for new customers are guided by the predicted CATEs and a predefined decision rule, such as targeting only those whose CATE exceeds the intervention cost. While statistically rigorous, this process prioritizes representativeness in the sampling step, rather than identifying and focusing on sampling the most consequential customers for the firm's profit-maximizing objective. Consequently, resources may be wasted on sampling customers whose inclusion has little impact on overall profitability, and with finite sample sizes targeting decisions may be suboptimal.

To illustrate this issue, consider a scenario where a firm aims to maximize profit by distributing a \$2 coupon through a targeted policy. The optimal strategy would involve targeting customers whose incremental spending due to the coupon (CATE) exceeds the coupon cost (i.e., \$2), thereby generating a positive incremental profit. In practice, however, a customer's true CATE is not observed directly, so firms must rely on predictions to guide their targeting decisions. These predictions, however, contain errors, which affect the firm's profitability in different ways depending on the customer. For illustration, let's consider three customers, as depicted in Figure 1.

For customer C, whose actual incremental spending is significantly above \$2 (or, more broadly, any customer whose true CATE is far from the decision threshold), moderate errors in CATE estimation are relatively inconsequential, as the firm can still make the correct targeting decision despite some degree of error. Conversely, for customers whose incremental spending is close to \$2 (i.e., whose true CATEs are near the decision threshold), even small prediction errors can lead to mistargeting. This scenario applies to customers A and B in Figure 1. Importantly, the impact of mistargeting either A or B differs in terms of profitability. Mistargeting customer A has minimal effect because the profit earned by the company from customer A remains nearly unchanged regardless of whether the customer is targeted. By contrast, mistargeting customer B leads to a notable profit loss, as their true incremental spending (CATE) is actually greater than \$2. In general, this highlights a key insight: when CATE estimation is used to deploy targeted policies, some customers warrant more precise attention — particularly those for whom prediction errors are most consequential for profitability. Intuitively, these are the customers with CATEs that are close to, but slightly deviate from, the decision threshold.

Figure 1 Differential Impact of Prediction Errors on Targeting Profitability



Note. This figure illustrates how prediction errors in CATE estimation impact targeting profitability differently. Customers A, B, and C have identical prediction errors. For customer C, whose true CATE is far from the threshold, the firm can still make the correct decision despite the error. For customer A, near the threshold, the error leads to mistargeting, but with minimal profit loss. In contrast, for customer B, whose CATE is slightly above the threshold, mistargeting due to the error causes a significant profit loss.

To address this issue, we propose a novel approach that aligns experimental design with the firm's profit-maximizing objectives. Unlike conventional methods, our approach prioritizes customers whose prediction errors in CATE estimation are most consequential to profitability, allowing firms to sample them intensively. Specifically, we introduce a sequential experimental design coupled with a novel sampling criterion — expected profit loss — that integrates the differential need for accuracy into the experimental design step. In particular, instead of randomly sampling all customers at once, we propose to split the predetermined experimental size into several smaller batches. The first batch includes a uniform sampling of customers similar to the traditional experimental design. Subsequent batches sequentially sample customers with the largest expected profit loss, as estimated from previous batches. We demonstrate that our expected profit loss sampling strategy aligns with the firm's profit-maximizing objective and introduce an expected profit loss estimation method leveraging the power of Bayesian inference for uncertainty quantification based on Causal Forest (Wager and Athey 2018, Athey et al. 2019b), a state-of-the-art non-parametric CATE model.

Through extensive simulation studies, we demonstrate the effectiveness of our proposed approach in identifying the most profitable targeting policy. Our method not only enhances the traditional test-then-learn practice but also outperforms leading adaptive experimental designs commonly referred to as 'test-and-learn' approaches. Specifically, we compare our method to the state-of-the-art adaptive experiment design for policy learning proposed by Kato et al. (2024), which sequentially assigns customers to treatment arms based on uncertainty rather than selectively sampling customers whose CATE estimation errors significantly affect targeting profitability. Our approach is particularly advantageous in cases where there are few customers near the decision threshold, as it employs an expected profit loss sampling criterion to effectively identify and intensively sample the most consequential customers. Notably, our method maintains its effectiveness even when the experimental sample is split into just two batches, with profit loss sampling only in the second stage. This two-stage design enables firms to shorten the experimentation period and streamline the process, which is especially useful when there is a delay between the intervention and the desired outcome (e.g., future sales).

We demonstrate the practical value of our proposed approach with two empirical applications. Using experimental data from a telecommunication company and a global coffeehouse chain (Starbucks), we show that our approach can generate more profitable targeting policies compared to current practices, particularly in scenarios where few customers are near the decision threshold. Such scenarios emerge from a misalignment between intervention costs and the central tendency of customer responsiveness, especially when: (1) the intervention is detrimental to a substantial number of customers (e.g., Ascarza et al. 2016), causing the distribution of treatment effects to cluster around negative values, whereas the cost of intervention (decision threshold) is always zero or positive; (2) the intervention carries the risk of cannibalizing revenues/profits that would have occurred without the intervention. For instance, providing free goods or monetary incentives may lead to negative incremental spending of targeted customers because some of them would have spent more without the offer. This dynamic is common in campaigns involving free credits, discounts, or coupons (e.g., Anderson and Simester 2004, Ailawadi et al. 2007, Ascarza 2018, Yoganarasimhan et al. 2023, Yang et al. 2023), as demonstrated in our empirical applications, where such promotions led to predominantly negative CATE values; or (3) the intervention itself is costly (e.g., phone calls, mailings) while customer responsiveness remains low (e.g., Lemmens and Gupta 2020, Simester et al. 2022), resulting in mostly-zero treatment effects and a positive decision threshold.

Collectively, our findings underscore the importance and value of a policy-aware experimentation approach that integrates the firm's business objectives directly into the experimental design.

The contributions of this research are three-fold. Methodologically, we propose a novel sampling criterion, the expected profit loss sampling strategy. The proposed approach sequentially samples customers for the experiment to maximize the firm's targeting profitability. We show that our approach effectively identifies the most consequential customers for selective sampling. Substantively, we demonstrate that firms can enhance their targeting performance by integrating their business objectives into the experimental design, and highlight the significance and benefits of a policy-aware experimentation approach. Managerially, we provide firms with a simple and efficient solution to enhance their targeting policies without increasing the sample size. This solution is achievable with only two stages, making it easy to implement in a relatively short experimentation period.

The paper is organized as follows. Section 2 provides a brief review of the literature. Section 3 introduces the policy-aware experimental design coupled with the expected profit loss sampling strategy. Section 4 presents the simulation results demonstrating the superiority of our proposed approach in enhancing targeting performance. Section 5 further validates the effectiveness of our approach through two empirical applications. Finally, Section 6 concludes the paper and discusses directions for future research.

2. Related Literature

Our paper contributes to several streams of literature. First, we add to the literature on targeting and personalization in marketing (e.g., Ascarza 2018, Hitsch and Misra 2018, Lemmens and Gupta 2020, Simester et al. 2020, Ellickson et al. 2022, Yang et al. 2023, Huang and Ascarza 2024), by identifying limitations in current practices and proposing a novel solution for experimentation sampling. Specifically, we highlight the importance of bringing the firms' business objectives into the design of the experiment. We demonstrate that by integrating firms' business objectives into the experimental design, the profitability of targeting policies can be effectively and efficiently enhanced without increasing the sample size.

Second, our work also relates to the experimental design literature (e.g., Blattberg 1979, Ginter et al. 1981, Feit and Berman 2019, Simester et al. 2022, Hu et al. 2024). This body of literature typically focuses on either determining the optimal experimental size to achieve a specific objective (Blattberg 1979, Ginter et al. 1981, Feit and Berman 2019, Simester et al. 2022), or devising the optimal sampling strategy across different pre-determined customer segments with the objective of minimizing the regret of the estimated policy (Hu et al. 2024). Our work contributes to this literature by introducing a novel sampling criterion that enhances targeting effectiveness by aligning firms' business objectives with the experimental design. Notably, our proposed method does not rely on any pre-segmentation of customers as in Hu et al. (2024), making it applicable in various business contexts where companies are uncertain about the true and relevant segmentation of customers.

Our research contributes to the expanding literature on adaptive experiments. This literature can be categorized into two distinct streams: the first stream is the bandit literature, which addresses the balance between exploration and exploitation in experimentation to reduce experimental costs (e.g., Hauser et al. 2009, Schwartz et al. 2017, Misra et al. 2019, Caria et al. 2020, Aramayo et al. 2022, Athey et al. 2022, Jain et al. 2024, Waisman et al. 2024). The second stream focuses on best arm identification (BAI), which aims to design treatment assignment rules to identify the most effective treatment arm with the minimum sample size (e.g., Bubeck et al. 2010, Chick and Frazier 2012, Grover et al. 2018, Jedra and Proutiere 2020, Kasy and Sautmann 2021, Carranza et al. 2023, Kato et al. 2024). Our work distinguishes itself from this literature by fundamentally shifting the focus from assigning customers to treatment arms to strategically sampling customers for the experiment. Specifically, instead of determining which treatment arm should be assigned to a (specific) customer (e.g., whether the customer should be treated or not), our method focuses on selecting which customers should be included in the experiment (and then randomly assigned to the treatment or control conditions) in order to minimize the impact of prediction error in CATE estimation on targeting profitability. Moreover, we address a common limitation of adaptive designs by demonstrating the effectiveness of our method when implemented in a simple two-stage design. This simplified approach is appealing to business practitioners as it avoids the technical and practical complexities associated with fully adaptive designs.

Third, this research is closely related to the active learning literature. A significant portion of this stream of research focuses on developing acquisition functions for sample queries to reduce the cost of data collection for supervised learning tasks (e.g., Fu et al. 2013, Wang and Ye 2015, Cardoso et al. 2017). Recently, some research has shifted their focus to identifying the most efficient selective sampling strategy for improving individual treatment effect estimation accuracy(Puha et al. 2020, Jesson et al. 2022) and decision-making tasks (Sundin et al. 2019, Filstroff et al. 2021, Liu et al. 2023). We build on this literature and develop a new acquisition function for selective sampling — expected profit loss — that incorporates the firm's profit-maximizing business objective into data acquisition by selectively sampling the customers whose errors in CATE estimation are most consequential for targeting profitability, thereby enhancing the performance of targeting policies estimated from the data.

Finally, our work relates to the emerging literature on decision-aware learning (e.g. Wilder et al. 2019, Wang et al. 2020, Kotary et al. 2021, Elmachtoub and Grigas 2022, Kallus and Mao 2023), which aims to align the prediction algorithm with the downstream optimization problem under the predict-then-optimize framework. While previous literature in various applications integrated decision objectives into prediction losses to enhance decision quality (e.g., (e.g. Lemmens and Gupta 2020, Chung et al. 2022)), we focus on the earlier experimentation phase. Furthermore, while previous research mostly focused on prediction tasks with observed outcomes, we focus on cases where the outcome of interest (i.e., CATE) is unobserved. Our proposed framework allows firms to improve their decision quality in various marketing contexts involving counterfactual inference of customers' responses to interventions.

3. Methodology

In this section, we first formulate the firm's targeting problem using the Neyman-Rubin potential outcomes model and characterize the firm's optimal sampling strategy to align with its profit-maximizing objective. Next, we address the practical challenges of implementing this sampling strategy and propose a sequential policy-aware experimental design with expected profit loss sampling as our solution. Additionally, we introduce an expected profit loss estimation strategy based on the Causal Forest model (Wager and Athey 2018, Athey et al. 2019b). This approach allows the firm to accurately identify consequential customers and sample them intensively without prior information about customers' responses to the intervention.

3.1. Problem Formulation

We consider a scenario where a firm aims to improve a specific outcome (Y_i) for each customer i through a marketing intervention with two treatment conditions ($W_i \in 0, 1$). The intervention incurs a cost $c(\cdot)$, which may vary across different customers (e.g., discount offers). Under this binary treatment, there are two potential outcomes: $Y_i(1)$ if the customer receives the intervention, and $Y_i(0)$ if the customer does not.

For each customer i, the firm observes a vector of customer characteristics $\mathbf{X}_i$ (e.g., demographics, past purchasing behavior) which may moderate the customer's response to the intervention. These characteristics are also known as pre-treatment covariates in the causal inference literature or targeting variables in marketing practice. The ultimate goal of the firm is to develop a targeting policy that, based on customer-observed covariates $\mathbf{X}_i$, assigns each customer i to the treatment condition that maximizes their profitability. In other words, the objective is to design the optimal targeting decision $\pi^*(\cdot) : \mathbf{X} \rightarrow \{0, 1\}$ such that

$$
\begin{aligned}
\pi^*(\mathbf{X}_i) &= \arg\max_{\pi} \mathbb{E}[\underbrace{Y_i(0) \cdot (1 - \pi(\mathbf{X}_i))}_{\text{profit w/o treatment}} + \underbrace{(Y_i(1) - c(\mathbf{X}_i)) \cdot \pi(\mathbf{X}_i)}_{\text{profit w/ treatment}} | \mathbf{X}_i] \\
&= \arg\max_{\pi} \mathbb{E}[Y_i(0) | \mathbf{X}_i] + (\underbrace{\mathbb{E}[(Y_i(1) - Y_i(0)) | \mathbf{X}_i]}_{\text{CATE}} - c(\mathbf{X}_i)) \cdot \pi(\mathbf{X}_i) \\
&= \arg\max_{\pi} \mathbb{E}[Y_i(0) | \mathbf{X}_i] + \underbrace{(\tau(\mathbf{X}_i) - c(\mathbf{X}_i))}_{\text{incremental profit}} \cdot \pi(\mathbf{X}_i)
\end{aligned}
\tag{1}
$$

where

$$\tau(\mathbf{X}_i) = \mathbb{E}[Y_i(1) - Y_i(0)|\mathbf{X}_i].$$

is the differential impact of the marketing intervention W_i conditioned on the pre-treatment covariates $\mathbf{X}_i$, which is also referred to as the conditional average treatment effect (CATE).

The expressions in (1) shows that the only determining factor of the firm's optimal targeting decision $\pi^*(\mathbf{X}_i)$ is the incremental profit of the intervention, defined as the difference between the CATE $\tau(\mathbf{X}_i)$ and the intervention cost $c(\mathbf{X}_i)$. Therefore, if the firm knows the true CATE, the optimal decision is to target the customers with a positive incremental profit such that

$$\pi^*(\mathbf{X}_i) = \mathbf{1}\{\tau(\mathbf{X}_i) > c(\mathbf{X}_i)\}. \tag{2}$$

However, in practice, the true CATE is never observed because for each customer, we either observe $Y_i(1)$ or $Y_i(0)$ due to the fundamental problem of causal inference (Holland 1986). Therefore, firms leverage experimental data D_e to estimate the CATEs, $\hat{\tau}(\mathbf{X}_i)$, and develop targeting policies, $\hat{\pi}(\mathbf{X}_i)$, based on these predictions; that is,

$$\hat{\pi}(\mathbf{X}_i) = \mathbf{1}\{\hat{\tau}(\mathbf{X}_i) > c(\mathbf{X}_i)\}. \tag{3}$$

We highlight three key aspects from the equations above. First, as presented in (3), the firm's targeting decision relies on the CATEs *estimated* from the experimental data, D_e, which inherently carry prediction errors. Therefore, the firm's decision to target a customer will be optimal *only* if their estimated CATE results in a targeting decision that matches the one based on their true CATE—i.e., when $\pi^*(\mathbf{X}_i) = \hat{\pi}(\mathbf{X}_i)$.

Second, because the targeting decision $\pi(\cdot)$ is non-linear (i.e., only targeting customers whose expected CATE exceeds the intervention cost, which can be interpreted as the decision threshold), the accuracy of $\hat{\tau}(\cdot)$ will differentially impact the optimality of targeting outcomes. For some customers, moderate prediction errors in τ will be inconsequential, as the firm's decisions remain unchanged despite the errors. Conversely, there are cases where even small prediction errors in CATE estimation can result in incorrect targeting decisions.

Third, as highlighted in (1), the profit loss incurred from mistargeting depends on the incremental profit for each customer, measured as the difference between their true CATE and the cost of the intervention. Thus, mistargeting customers with high incremental profit may be more consequential for the firm.

Crucially, since the firm controls the collection of the experimental data, D_e, it can increase the profitability of its targeting policies by employing a sampling procedure designed to minimize the impact of prediction errors. This "policy-aware" experimental design involves focusing on increasing the experimental samples around "consequential" customers—those for whom errors in CATE estimation are most consequential for profitability.

In the remainder of this section, we characterize the firm's optimal sampling strategy by identifying these consequential customers and demonstrating how this approach can achieve the firm's profit-maximizing objective.

3.2. Optimal Sampling Strategy

We propose a sampling strategy that identifies the most consequential customers based on their *expected profit loss*. Specifically, we first establish that customers consequential for firms aiming for a profit-maximizing targeting policy are those with high expected profit loss. We then mathematically demonstrate the effectiveness of the proposed expected profit loss sampling strategy in achieving the firm's profit-maximizing objective.

3.2.1. Characterizing Consequential Customers Consider a customer whose true CATE exceeds the decision threshold by a large amount $M > 0$, i.e., $\tau(\mathbf{X}_i) - c(\mathbf{X}_i) > M$ (see Figure 2a). The optimal decision for the firm would be to target this customer, as his/her treatment effect exceeds the cost, bringing an incremental profit of $\tau(\mathbf{X}_i) - c(\mathbf{X}_i)$. However, the firm does not know the true CATE, but only its prediction $\hat{\tau}(\mathbf{X}_i) = \tau(\mathbf{X}_i) - \varepsilon$, where ε captures the estimation error. In this case, the firm can still make the correct targeting decision as long as the prediction error ε is smaller than M.

Conversely, if the customer's true CATE exceeds the decision threshold by only a small amount $m > 0$ (see Figure 2b), even small estimation errors greater than m will lead to incorrect targeting decisions.[1]

Figure 2 illustrates how the same prediction error ε can lead to very different outcomes in the two scenarios. In particular, the firm can still make the correct targeting decision $\hat{\pi}(\cdot) = \pi^*(\cdot)$ if the customer's true CATE $\tau(\cdot)$ deviates significantly from the decision threshold. Conversely, customers whose true CATE $\tau(\cdot)$ is near the threshold are likely to be mistargeted.

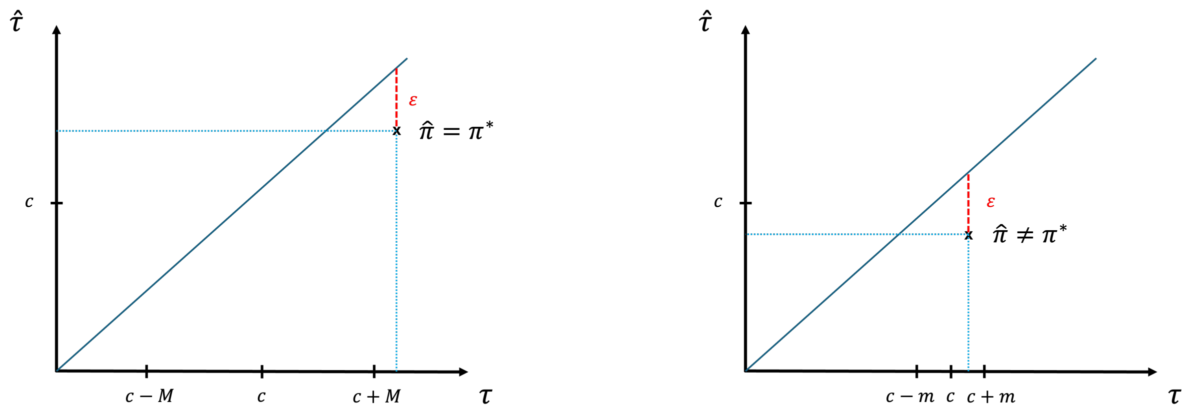
Following this characterization, one might conclude that the latter type of customers are the most consequential for the firm. However, mistargeting customers with CATEs directly around the decision threshold might not significantly impact profitability. For these customers, the true CATE is almost equal to the cost of targeting, resulting in minimal profit loss from incorrect targeting (see (1)). Therefore, the customers who are most crucial and deserve more attention are those whose CATEs are sufficiently but not very close to the decision threshold.

In summary, the consequential customers for the firm, who warrant more attention, are those with a high mistargeting probability $\mathbb{P}\left(\hat{\pi}(\mathbf{X}i) \neq \pi^*(\mathbf{X}i)\right) = \mathbb{E}\left[\mathbf{1}\{\hat{\pi}(\mathbf{X}i) \neq \pi^*(\mathbf{X}i)\}\right]$ (i.e., those whose

[1] Note that these arguments also hold for customers whose true CATE is smaller than the threshold.

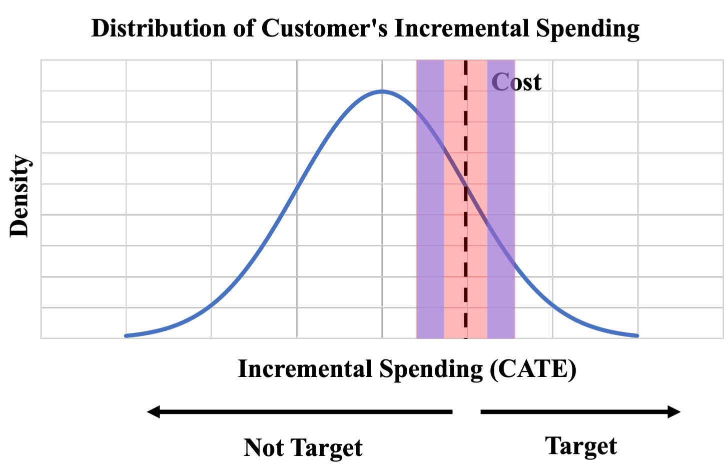
Figure 2 **Differential Impact of Prediction Error ε on Targeting Accuracy**

(a) $\tau(\mathbf{X}_i) - c(\mathbf{X}_i) > M$ (b) $\tau(\mathbf{X}_i) - c(\mathbf{X}_i) < m$



CATEs are near the decision threshold) but at the same time may lead to a relatively high associated profit loss $|\tau(\mathbf{X}i) - c(\mathbf{X}i)|$ (i.e., those whose CATEs is not very close to the threshold). Figure 3 provides an illustration of these consequential customers. Based on this insight, we next introduce the expected profit loss sampling strategy and demonstrate how it aligns with the firm's profit-maximizing objective.

Figure 3 **Illustration of Consequential Customers**



Note. Customers in the clear region have CATEs that deviate significantly from the decision threshold, allowing for moderate prediction errors without compromising targeting decisions. Customers in the red region have CATEs directly around the threshold, but mistargeting them has minimal impact on the firm's profit. However, customers in the purple region, whose CATEs are close to but slightly deviate from the threshold, warrant the most attention since prediction errors in this group can have a more significant effect on profitability.

3.2.2. The Expected Profit Loss Sampling Strategy We define expected profit loss as

$$\ell(\mathbf{X}_i) = \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \mathbf{1}\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\}\right], \tag{4}$$

which multiplies the two key quantities that determine a "consequential" customer — mistargeting probability and profit loss. The following proposition demonstrates that minimizing the prediction error in CATE estimation for customers with high expected profit loss aligns with the firm's profit-maximizing objective.

Proposition 1 **(Optimal Sampling Strategy)**. *To maximize the expected profit of targeting policy $\hat{\pi}(\cdot)$ generated from $\mathcal{D}_e$, the firm should focus on minimizing the absolute prediction error of $\hat{\tau}(\cdot)$ for the customers with higher expected profit loss $\ell(\cdot)$:*

$$\ell(\mathbf{X}_i) = \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \mathbf{1}\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\}\right].$$

Proof: See Web Appendix B.

Proposition 1 shows that by prioritizing the minimization of prediction error in CATE estimation for customers with high expected profit loss, the firm can identify the targeting policy that yields the highest expected profit. Achieving this requires the firm to gather more data on these customers. This proposition thus provides the mathematical foundation for our expected profit loss sampling strategy and demonstrates how the proposed criterion successfully integrates the firm's business objective into the experimental design.

3.2.3. Practical Challenges Although the expected profit loss sampling strategy theoretically aligns the firm's profit-maximizing objective with the experimental design, its implementation presents several challenges. The primary difficulty stems from the firm's limited knowledge of customers' heterogeneous responses to the intervention ($\tau(\mathbf{X}_i)$). Specifically, to calculate the expected profit loss sampling criterion, the firm must first estimate the CATEs and the corresponding targeting decisions for each customer. However, firms lack knowledge of customers' responses to the intervention. Thus, to initiate the process of policy learning, the firm needs to conduct an experiment to estimate $\tau(\mathbf{X}_i)$.

Additionally, even if the firm acquires the experimental data needed for CATE estimation, implementing the expected profit loss sampling strategy remains challenging due to the fundamental problem of causal inference. In particular, calculating the expected profit loss requires the firm to know the true CATEs of each customer. However, each customer can only be assigned to one of the two treatment conditions at a time, making the true CATEs unobserved by the firm. Therefore, addressing these challenges requires a solution that enables the firm to collect data on customers' responses to the intervention and accurately estimate the expected profit loss without knowing the true CATEs.

3.3. A Sequential Experimental Design with Expected Profit Loss Sampling

To tackle these challenges, we propose a sequential experimental design combined with a novel estimation method to compute expected profit loss sampling. Specifically, the sequential experimental design enables the firm to *strategically collect* experimental samples based on data gathered in earlier batches, facilitating more accurate CATE estimation for consequential customers. This is in contrast to the current practice of the "test-then-learn" approach (e.g. Ascarza 2018, Yoganarasimhan et al. 2023, Huang and Ascarza 2024) that samples all customers with equal probability.

To guide the sampling mechanism, we introduce an expected profit loss estimation strategy utilizing Causal Forest (Wager and Athey 2018, Athey et al. 2019b). This estimation method allows the firm to identify customers with high expected profit loss nonparametrically without any need for pre-segmenting customers.

Before getting into the proposed sampling approach, we make the following assumptions:

ASSUMPTION 1. *(Stable Unit Treatment Value Assumption, SUTVA) The potential outcomes for customer i in response to treatment W_i are independent of the treatment assignment of other customers i'. Formally,*

$$Y_i(\mathbf{W}) = Y_i(W_i)$$

.

ASSUMPTION 2. *(Stability) The potential outcomes for customer i in response to treatment W_i does not change overtime, meaning that for all $t \neq t'$*

$$\mathbb{E}[Y_i^t(W_i)|\mathbf{X}_i] = \mathbb{E}[Y_i^{t'}(W_i)|\mathbf{X}_i]$$

where $Y_i^t(W_i), Y_i^{t'}(W_i)$ denote the potential outcomes at time t and t' respectively.

The first assumption ensures that there is no interference between different customers, meaning that the treatment assignment for one customer does not influence another customer's behavior. The second assumption guarantees that customers' responses to the treatments remain consistent over time during the experimentation period and are representative of the targeting population. Consequently, the characteristics of customers with high expected profit losses remain unchanged throughout the experimental period and are indicative to those in the targeting population.

Given a firm's customer base $\mathcal{I}$ and a pre-determined experimental size N_e (with $N_e < \text{size}(\mathcal{I})$), we choose B as the number of batches for the sequential experimentation procedure.[2] For the first

[2] We focus exclusively on scenarios where firms have a fixed customer base to sample from, excluding cases where customers arrive sequentially and unpredictably. This is a common situation in many marketing applications, such as customer retention (e.g. Ascarza 2018, Lemmens and Gupta 2020, Yang et al. 2023) and catalog mailing campaign (e.g. Hitsch and Misra 2018, Simester et al. 2020).

batch ($b = 1$), we randomly sample $\frac{N_e}{B}$ customers from the customer base $\mathcal{I}$ and randomly assign them to the two treatment conditions $W_i \in \{0, 1\}$. Once the outcomes of the customers in the first batch are observed, we predict the expected profit loss for the remaining unsampled customers using the data collected from the first batch D_e^1. In the second batch ($b = 2$), we then select $\frac{N_e}{B}$ customers with the highest estimated expected profit loss from those previously unsampled, $\mathcal{I} \backslash D_e^1$, and randomly assign them to the two treatment arms. This process is repeated until the final batch ($b = B$).[3]

Note that treatment allocation is random in all batches; the only difference from current practice is the types of customers *chosen to be sampled* for the experimentation, where we prioritize those with higher expected profit loss. This is also in contrast with commonly-used bandit/BAI sampling strategies such as Thomson sampling (e.g. Schwartz et al. 2017, Jain et al. 2024) and upper confidence bound (UCB) algorithm (e.g. Misra et al. 2019) that allocate units to specific treatment arms. In our proposed approach, the treatment allocation is always random.

Once the outcomes of the customers in the final batch are observed, we train a Causal Forest model using the data collected across all B batches, D_e. Given that each customer in the experiment is randomly assigned to one of the two treatment arms, the following two assumptions hold:

Assumption 3. *(Unconfoundedness) For every customer i, the treatment assignment W_i is independent of any unobserved factors, i.e.,*

$$(Y_i(0), Y_i(1)) \perp\!\!\!\perp W_i | \mathbf{X}_i$$

.

Assumption 4. *(Overlap) For every customer i, the propensity of being targeted is neither* 0 *or* 1, *i,e.,*

$$0 < Pr(W_i = 1 | \mathbf{X}_i = x) < 1, \forall x$$

.

Consequently, the CATE prediction for each customer i generated by the CATE model trained on the experimental data collected through this sequential design is consistent (i.e., $\mathbb{E}[\hat{\tau}(\mathbf{X}_i)] = \tau(\mathbf{X}_i)$).[4] We leverage these CATE predictions to derive the final targeting policy as described in Section 3.1. The pseudo-code for the solution is outlined in Algorithm 1.

[3] Selecting the number of batches B requires balancing targeting precision against feasibility — additional batches improve identification of consequential customers but lengthen experimental duration. Therefore, We recommend setting B as the maximum number of acceptable batches determined by practical constraints such as experiment timeline.

[4] Note that the ATE estimated by our approach would be inconsistent since the experimental data is unrepresentative of the population.

Algorithm 1 The Sequential Experimental Design with Expected Profit Loss Sampling

Customer Base: $\mathcal{I}$

Input: Experimental size $N_e \in \mathbb{N}$; Number of batches $B \in \mathbb{N}$

Output: Targeting Decision $\hat{\pi}$

for b = 1, 2, ..., B **do**

  if b=1 **then**

    Randomly sample $\frac{N_e}{B}$ customers from $\mathcal{I}$

    Randomly assign each sampled customer i to the two treatment arms $W_i \in \{0,1\}$

    Observe the outcomes of the sampled customers Y_i

  else

    For each unsampled customer $i \notin D_e^{b-1}$, estimate the expected profit loss $\hat{\ell}_b(\mathbf{X}_i)$ with the data collected in previous $b-1$ batches D_e^{b-1}

    Select $\frac{N_e}{B}$ customer with the highest expected profit loss $\hat{\ell}_b(\mathbf{X}_i)$

    Randomly assign each customer i in this batch to the two treatment arms $W_i \in \{0,1\}$

    Observe the outcomes of the sampled customers Y_i

  end if

end for

Estimate a Causal Forest model with the experimental data D_e^B

Derive the final targeting decision leveraging the CATE predictions:

$$\hat{\pi}(\cdot) = \mathbf{1}\{\hat{\tau}(\cdot) > c(\cdot)\}.$$

3.3.1. Estimating Expected Profit Loss A key component of our proposed solution is the estimation of expected profit loss, which enables the accurate identification of customers with the highest expected profit loss, even without direct observation of true CATEs. To address this challenge, we leverage the power of Bayesian inference for uncertainty quantification and propose an estimation method based on the Causal Forest model (Wager and Athey 2018, Athey et al. 2019b).

Specifically, prior to starting batch b, we estimate the CATE, $\hat{\tau}_b(\mathbf{X}i)$, and its posterior predictive distribution, $p_b(\tau(\mathbf{X}i) \mid D_e^{b-1})$, for each customer not sampled in the previous $b-1$ batches, where D_e^{b-1} represents the data collected up to batch $b-1$. This posterior predictive distribution, $p_b(\tau(\mathbf{X}_i) \mid D_e^{b-1})$, quantifies the uncertainty of the CATE estimates $\hat{\tau}(\mathbf{X}_i)$ derived from the collected data, enabling us to estimate the likelihood of deviations between the estimated targeting decision $\hat{\pi}(\mathbf{X}_i)$ and the optimal targeting decision, as well as the associated profit loss.

Using the estimated CATE, $\hat{\tau}_b(\mathbf{X}i)$, we determine the targeting decision, $\hat{\pi}_b(\mathbf{X}i)$, as $\hat{\pi}_b(\mathbf{X}i) = \mathbf{1}\{\hat{\tau}_b(\mathbf{X}i) > c(\mathbf{X}i)\}$. We then estimate the expected profit loss, $\hat{\ell}_b(\mathbf{X}_i)$, associated with the targeting decision by integrating all possible profit losses in the posterior predictive distribution, $p_b\left(\tau(\mathbf{X}_i)|D_e^{b-1}\right)$. That is,

$$\begin{aligned}\hat{\ell}_b(\mathbf{X}_i) &= \int \underbrace{0 \cdot \mathbf{1}\{\hat{\pi}_b(\mathbf{X}_i) = \pi(\mathbf{X}_i)\}}_{\text{profit loss with no deviation}} + \underbrace{|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|\mathbf{1}\{\hat{\pi}_b(\mathbf{X}_i) \neq \pi(\mathbf{X}_i)\}}_{\text{profit loss with deviation}} p_b(\tau(\mathbf{X}_i)|\mathcal{D}_e^{b-1})d\tau \\ &= \int \underbrace{|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|\mathbf{1}\{\hat{\pi}_b(\mathbf{X}_i) \neq \pi(\mathbf{X}_i)\}}_{\text{profit loss}} p_b(\tau(\mathbf{X}_i)|\mathcal{D}_e^{b-1})d\tau,\end{aligned} \tag{5}$$

where $\tau(\mathbf{X}_i)$ is a draw from the posterior predictive distribution and $\pi(\mathbf{X}_i) = \mathbf{1}\{\tau(\mathbf{X}_i) > c(\mathbf{X}_i)\}$. Intuitively, when $\hat{\tau}_b(\mathbf{X}i)$ is far from the decision threshold, only a small portion of the draws from its posterior predictive distribution are likely to produce a different targeting decision $\pi(\mathbf{X}i)$ from $\hat{\pi}_b(\mathbf{X}i)$, leading to a low expected profit loss estimate, $\hat{\ell}_b(\mathbf{X}i)$. In contrast, when $\hat{\tau}_b(\mathbf{X}i)$ is directly around the threshold, the associated profit loss $|\tau(\mathbf{X}i) - c(\mathbf{X}i)|$ is likely small for most draws, preventing a high expected profit loss estimate. Finally, when $\hat{\tau}_b(\mathbf{X}i)$ is close but slightly deviates from the decision threshold, more draws from its posterior predictive distribution are likely to yield different targeting decisions from $\hat{\pi}_b(\mathbf{X}i)$ with higher profit losses. This scenario results in the highest expected profit loss estimate, prompting the algorithm to prioritize sampling these customers.

A crucial step in our estimation strategy is deriving the posterior predictive distribution, $p_b(\tau(\mathbf{X}_i)|\mathcal{D}_e^{b-1})$. To avoid restrictive parametric assumptions about the CATE distribution, we leverage the Causal Forest model (Wager and Athey 2018, Athey et al. 2019b) and approximate the posterior predictive distribution by bootstrapped samples. Specifically, for each batch b, we first train a Causal Forest model with J trees on the data collected in previous $b-1$ batches, D_e^{b-1}, and obtain $\hat{\tau}_b(\mathbf{X}_i)$. Since each tree in the Causal Forest model is estimated on different bootstrapped samples, we utilize the fact that the bootstrap distribution approximates the posterior distribution with a noninformative prior (Hastie et al. 2009) and treat the prediction of each single tree j in the forest, $\tilde{\tau}_b^j(\mathbf{X}_i)$, as a draw from the posterior predictive distribution, $p_b(\tau(\mathbf{X}_i)|\mathcal{D}_e^{b-1})$.[5]

Next, we determine the optimal targeting decision $\tilde{\pi}_b^j(\mathbf{X}_i)\} = \mathbf{1}\{\tilde{\tau}_b^j(\mathbf{X}_i) > c(\mathbf{X}_i)\}$ and calculate the profit loss $|\tilde{\tau}_b^j(\mathbf{X}_i) - c(\mathbf{X}_i)|\mathbf{1}\{\hat{\pi}_b(\mathbf{X}_i) \neq \tilde{\pi}_b^j(\mathbf{X}_i)\}$ associated with each draw $\tilde{\tau}_b^j(\mathbf{X}_i)$. Finally, we

[5] We chose Causal Forest as our estimation method primarily due to its practical implementation advantages. Firms can readily implement our expected profit loss estimation approach using existing packages like `grf` in R and `econML` in Python, enhancing the practical appeal of our method. While our approach is demonstrated using Causal Forest, practitioners can follow a similar procedure using alternative methods such as Bayesian nonparametric models that allows for uncertainty quantification.

approximate the integral in Equation 5 by averaging the profit losses associated with the draws $\tilde{\tau}_b^j(\mathbf{X}_i)$ across all J trees to obtain the expected profit loss estimate, $\hat{\ell}_b(\mathbf{X}_i)$. The pseudo-code for the estimation algorithm is outlined in Algorithm 2.

Algorithm 2 Estimating Expected Profit Loss

Input: Number of trees $J \in \mathbb{N}$

Output: $\hat{\ell}_b(\mathbf{X}_i)$

Data: D_e^{b-1}

Train a Causal Forest model on D_e^{b-1} and obtain $\hat{\tau}_b(\mathbf{X}_i)$.

for j = 1, 2, ..., J **do**

Determine the optimal targeting decision associated with the prediction $\tilde{\tau}_b^j(\mathbf{X}_i)$:

$$\tilde{\pi}_b^j(\mathbf{X}_i)\} = \mathbf{1}\{\tilde{\tau}_b^j(\mathbf{X}_i) > c(\mathbf{X}_i)\}.$$

Calculate the profit loss associated with the prediction $\tilde{\tau}_b^j(\mathbf{X}_i)$ by

$$\hat{\ell}_b^j(\mathbf{X}_i) = |\tilde{\tau}_b^j(\mathbf{X}_i) - c(\mathbf{X}_i)|\mathbf{1}\{\hat{\pi}_b(\mathbf{X}_i) \neq \tilde{\pi}_b^j(\mathbf{X}_i)\}.$$

end for

Calculate the expected profit loss estimate by averaging the profit losses associated with the draws across all J trees:

$$\hat{\ell}_b(\mathbf{X}_i) = \frac{1}{J}\sum_{j=1}^{J} \hat{\ell}_b^j(\mathbf{X}_i).$$

return $\hat{\ell}_b(\mathbf{X}_i)$

3.4. Addressing Delayed Feedbacks: A Simplified Two-Stage Design

A key challenge in implementing our sequential approach arises from the time lag between marketing interventions and observed customer responses, a limitation that becomes particularly salient when using a larger number of batches.[6] For example, a firm aiming to boost customers' next-week spending with a coupon can only observe the desired outcome a week later. This delay complicates the adoption of a fully adaptive design, as the experimentation process would become too lengthy and impractical. Additionally, technical complexities and engineering costs associated with the implementation of an adaptive experiment (Hadad et al. 2021), may prevent the firm from adopting a fully adaptive design in practice.

[6] One may argue that this can be solved by using intermediate outcomes as surrogates for delayed feedback (Athey et al. 2019a, Yang et al. 2023, Huang and Ascarza 2024). However, these intermediate outcomes may also require significant time to observe post-intervention.

To mitigate these concerns, we propose a simplified two-stage version of our approach that maintains effectiveness while reducing implementation complexity. In this simplified design, we divide the experimental sample into two (potentially unequal) groups, applying our expected profit loss sampling only in the second stage. This modification substantially reduces both the experimentation duration and operational complexity, making it more appealing to practitioners.

Unlike a multi-stage sequential design that allows for iterative refinement of expected profit loss estimates and consequential customer identification, a two-stage design offers only one opportunity for strategic sampling. This creates a critical trade-off: we need enough customers in the first stage to generate accurate expected profit loss estimates, while reserving sufficient customers for strategic sampling in the second stage. To provide practical guidance on this balance, we examine what portion of the total customer sample should be allocated to the first stage through both simulation studies and empirical applications in the following sections.

4. Simulation

We perform a series of simulation studies to achieve three main objectives. First, we examine the performance of our method in learning the profit-maximizing targeting policy, particularly in situations where the number of customers around the decision threshold is limited. Second, we explore the efficacy of our approach in accurately identifying customers whose prediction errors in CATE estimation could significantly impact the firm's profitability, as illustrated in Section 3.2.1. Third, we show how our approach can be simplified into a two-stage design while still maintaining its effectiveness, thereby addressing the practical challenges associated with a fully adaptive design as discussed in Section 3.4.

4.1. Simulation Setup

We consider the scenario where a firm aims to develop a targeting policy that maximizes the profitability of a marketing intervention. As it is commonly implemented in practice, the firm first learns the targeting policy through experimentation on a subset of customers, and subsequently implements the learned policy in the remaining (and/or future) customers. We assume that the impact of the (binary) intervention on customers is heterogeneous and follows a normal distribution centered at 1, meaning on average, the intervention has a positive effect on consumers.[7]

Specifically, we generate a customer base $\mathcal{I}$ and an evaluation set D_{eval} with a binary treatment $W_i \in \{0, 1\}$ according to the following data generating process:

$$Y_i = \tau(X_i) \cdot W_i + X_{i4} \cdot X_{i5} + e_i$$

[7] We also explore alternative scenarios in which the CATE distribution is bimodal, featuring two segments of equal or unequal proportions. The results are available in Web Appendix C.2.

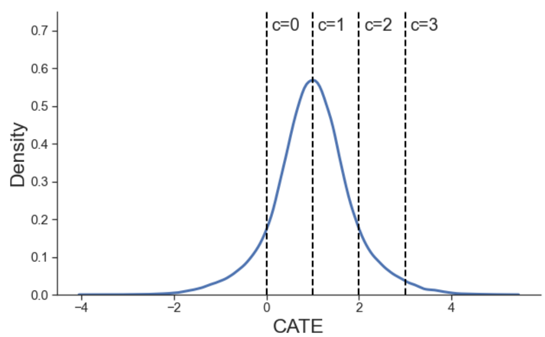
where

$$\tau(X_i) = X_{i1} \cdot X_{i2} + 0.5 \cdot X_{i3} + 1,$$
$$e_i \sim \mathcal{N}(0,1),$$
$$X_{ij} \sim \mathcal{N}(0,1), \quad j \in \{1,3,5\},$$
$$X_{ij} \sim \text{Bernoulli}(0.5), \quad j \in \{2,4\}.$$

This data generating process follows Assumptions 1 and 2.

We assume that the marketing intervention incurs a cost, c, which represents the decision threshold. In our simulation studies, we hold the (simulated) CATE distribution constant, with a mode and median equal to 1, and vary the intervention cost with $c \in \{0,1,2,3\}$ to examine our approach's performance across different scenarios (see Figure 4 for an illustration).[8]

Figure 4 CATE Distribution and Intervention Costs



Note. Each dashed line corresponds to a different intervention cost c. $c = 1$ corresponds to the mode of the CATE distribution. $c = 3$ represents the maximum deviation from the mode. $c = 0$ and $c = 2$ are symmetrically positioned around the mode.

Specifically, $c = 0$ represents scenarios with minimal intervention costs and positive impact; while such scenarios are rare in practice, an example might be an email campaign where the objective is simply for recipients to open the email. The case $c = 1$ represents scenarios where customers are predominantly clustered around the decision threshold. By contrast, $c = 3$ describes situations where only a few customers are near the decision threshold. This latter case is common in practice, either because treatment effects are generally low across customers or because intervention costs are prohibitively high (e.g., Ascarza et al. 2016, Lemmens and Gupta 2020).

Additionally, because $c = 0$ and $c = 2$ are symmetrically positioned around the mode of the CATE distribution (located at 1, making their distance to the mode equal), this setup allows us

[8] Alternatively, we could vary the CATE distribution itself, but because only the difference between CATE and cost matters for the targeting decision, both approaches yield same insights.

to investigate whether our approach's performance is influenced solely by the distance from the decision threshold to the mode, or also by the direction relative to the mode (with $c = 0$ below and $c = 2$ above). We also vary the experimental size ($N_e \in \{1\text{k}, 5\text{k}, 10\text{k}, 20\text{k}, 30\text{k}\}$) to assess sample size efficiency.

4.2. Experimental Designs for Comparison

4.2.1. Policy-Aware We examine two variants of our proposed policy-aware approach: one employing multiple batches, with each batch comprising the same number of customers (in this case, 200 customers), and another utilizing only two batches of varying size.[9] In both cases, the total experimental size remains constant; the only difference is the number of batches drawn, which has implications for the number of times the researchers need to observe consumer outcomes and the number of times the expected profit loss sampling strategy is employed. This comparison enables us to evaluate the performance differences between a more complex adaptive design and a simplified two-stage design, thereby elucidating the potential trade-offs between implementation complexity and targeting efficacy. For the two-stage design, we investigate various proportions of customers r to be sampled in the first stage ($r \in \{0.5, 0.7, 0.9\}$), aiming to examine the optimal two-step configuration for different scenarios.

4.2.2. Benchmarks We compare our approach with two alternative experimental designs. The first is the test-then-learn approach (`Default`) commonly employed in practice (e.g., Ascarza 2018, Yang et al. 2023, Huang and Ascarza 2024). This method involves randomly sampling customers with equal probability for the experiment and assigning them to different treatment arms at random. In our simulation, we assign the sampled customers to the two treatment arms with a probability of 0.5.

We also compare our approach to the state-of-the-art adaptive experimental design (`Adaptive`) proposed by Kato et al. (2024). While most adaptive designs in the literature focus on different objectives (e.g., balancing exploration and exploitation) or require strong parametric assumptions, Kato et al. (2024) aims to estimate the most accurate targeting policy with limited parametric assumptions. This alignment in objectives and assumptions makes it an ideal method for comparison with our approach. Note that, as designed, this adaptive approach emphasizes treatment assignment decisions for sampled customers rather than determining which customers to include

[9] For the multi-stage design, we follow a similar procedure as Waisman et al. (2024) by using equal-sized batches throughout the experiment.

in the experiment. Specifically, the Kato et al. (2024) approach assigns the customers sampled in batch b to different treatment arms based on the following rule:

$$P_b(W_i = 0|\mathbf{X}_i = x) = \frac{\sigma_b^0(x)}{\sigma_b^0(x) + \sigma_b^1(x)}$$
$$P_b(W_i = 1|\mathbf{X}_i = x) = \frac{\sigma_b^1(x)}{\sigma_b^0(x) + \sigma_b^1(x)}$$

where $\sigma_b^w(x)$ denotes the standard deviation of the potential outcomes $Y_i(W_i = w)$ estimated from the previous $b-1$ batches.[10] Intuitively, this approach prioritizes assigning customers to the treatment arm with greater uncertainty in expected outcomes, allowing the firm to reduce uncertainty in customer responses and identify the most effective treatment for each customer more accurately. This is in contrast to our approach, where we selectively sample the customers whose prediction errors in CATE estimation have the most significant impact on the firm's profitability, and randomly assign them to different treatment conditions.

4.3. Evaluation Procedure

We evaluate the targeting performance of each of the learned policies in a validation set, simulating the company implementing the learned policy in the future. For each policy approach, we run 100 replications of the experimental design, calculate the optimal policy using the experimental data, calculate the profitability of the resulting policy, and report the mean of key metrics. Initially, we generate $N_{\mathcal{I}} = 100,000$ customers as our customer base, $\mathcal{I}$, and $N_{eval} = 10,000$ customers as the evaluation set, D_{eval}, which will be used across all bootstrap samples. In each replication, we sample N_e customers from the customer base, $\mathcal{I}$, for the experiment and estimate the CATE model based on the experimental data, D_e, for each approach. We then leverage the constructed CATE model to generate targeting decisions, $\hat{\pi}(\cdot)$, on the evaluation set, D_{eval}. Finally, we calculate the evaluation metric, the proportional profit gap associated with the targeting decisions, on the evaluation set using the following formula:

$$\text{PPG} = \frac{\overbrace{\sum_{i \in D_{eval}} (\tau(X_i) - c) \cdot \pi^*(X_i)}^{\text{incremental profit of } \pi^*(\cdot)} - \overbrace{\sum_{i \in D_{eval}} (\tau(X_i) - c) \cdot \hat{\pi}(X_i)}^{\text{incremental profit of } \hat{\pi}(\cdot)}}{\underbrace{\sum_{i \in D_{eval}} (\tau(X_i) - c) \cdot \pi^*(X_i)}_{\text{incremental profit of } \pi^*(\cdot)}}$$

where $\pi^*(\cdot)$ represents the true optimal targeting decision based on the true CATE $\tau(\cdot)$. Note that this evaluation metric quantifies the difference between the profit generated by the estimated

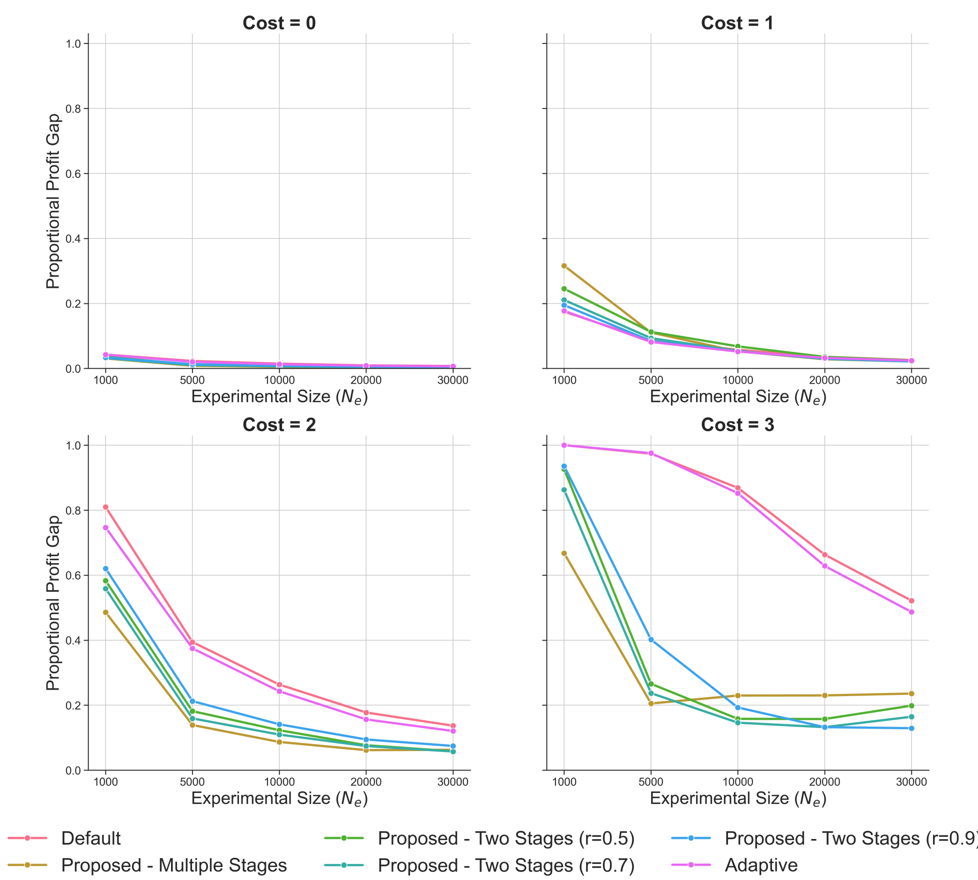
[10] To account for the adaptive nature of the experimental data, we follow Kato et al. (2024) and reweight the outcome Y_i for customer i sampled in batch b by $P_b(W_i|\mathbf{X}_i)$. See Web Appendix C.1 for implementation details.

targeting policies $\hat{\pi}(\cdot)$ and the true optimal policy $\pi^*(\cdot)$. Therefore, a smaller profit gap indicates that the targeting decision estimated from the experimental data D_e is more effective in finding the profit-maximizing targeting policy.

4.4. Results

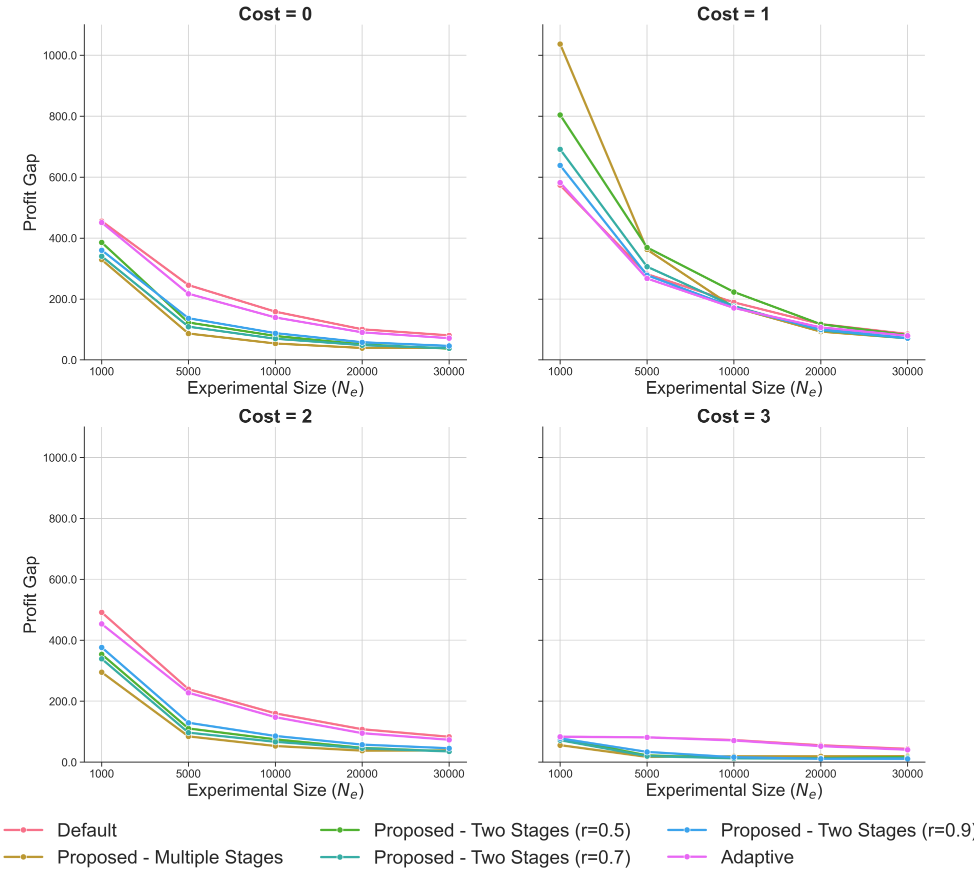
4.4.1. Profitability of Targeting Policies Figure 5 shows the proportional profit gaps of the targeting policies learned by different experimental designs across various intervention costs ($c \in \{0, 1, 2, 3\}$) and experimental sizes ($N_e \in \{1k, 5k, 10k, 20k, 30k\}$). To further examine the performance symmetry between $c = 0$ and $c = 2$, we also report the profit gap of the targeting policies in Figure 6, alongside the proportional profit gap. This approach helps us evaluate performance without the potential distortion that can arise from using different denominators in the proportional metric.

Figure 5 Proportional Profit Gaps of Different Experimental Designs



Note. We report the average value of the proportional profit gap across 100 replications. Each line corresponds to an experimental approach.

Figure 6 Profit Gaps of Different Experimental Designs



Note. We report the average value of the profit gap (i.e., the difference between the incremental profit of $\hat{\pi}(X_i)$ and that of the optimal policy) across 100 replications. Each line corresponds to an experimental approach.

The results reveal several interesting insights. First, when the cost deviates from the mode of the CATE distribution (i.e., $c \in \{0, 2, 3\}$), our approach consistently outperforms the two benchmarks, delivering more profitable targeting policies. Notably, the additional profits achieved by our approach (in absolute terms) are similar between $c = 0$ and $c = 2$, indicating that the benefit of our approach is driven by the distance between the decision threshold and the mode of the CATE distribution.[11] However, when the decision threshold aligns with the mode (i.e,. $c = 1$), our approach performs comparably to the benchmarks and may even underperform when the sample size is small ($N_e < 10000$). This occurs because, when the decision threshold is close to the mode of the CATE distribution, random sampling inherently focuses on the consequential customers effectively. In this scenario, the relatively imprecise expected profit loss estimate when sample sizes

[11] The difference in the proportional profit gap between $c = 0$ and $c = 2$ is primarily due to the disparity in their denominators. In particular, because $c = 0$ represents a less costly intervention, the incremental profit generated by the optimal policy is greater compared to $c = 2$. As a result, despite having similar absolute profit gap, the proportional profit gaps differ between the two scenarios.

are small may misguide the sampling process in our approach, resulting in less intensive sampling of consequential customers compared to random sampling.

Beyond providing numerical evidence of the proposed method's performance, these findings highlight the scenarios in which policy-aware experimentation offers the greatest benefit to firms. Specifically, when the number of customers near the decision threshold is limited, our approach enables firms to develop more profitable targeting policies compared to other designs. Such scenarios emerge from a misalignment between intervention costs and the central tendency of customer responsiveness, especially when: (1) the intervention negatively impacts a substantial portion of customers (e.g., Ascarza et al. 2016), causing the distribution of treatment effects to cluster around negative values, whereas the cost of intervention (decision threshold) is always zero or positive; (2) the intervention itself is costly (e.g., phone calls, mailings) while customer responsiveness remains low (e.g., Lemmens and Gupta 2020, Simester et al. 2022), resulting in mostly-zero treatment effects and a positive decision threshold; and (3) the intervention introduces a risk of cannibalizing profits that would have naturally occurred without intervention, such as when providing free goods or monetary incentives (e.g., Anderson and Simester 2004, Ailawadi et al. 2007, Ascarza 2018, Yoganarasimhan et al. 2023, Yang et al. 2023), leading to predominantly negative CATE values. The latter scenario is further explored in Section 5, where we analyze two empirical datasets featuring real-world marketing campaigns with those characteristics.[12]

Second, we emphasize the sample efficiency of our approach. Specifically, when the cost deviates from the mode of the CATE distribution (i.e., $c \in \{0, 2, 3\}$), our approach performs comparably to the two benchmarks with larger sample sizes. For example, when $c = 2$, the targeting policy estimated from a sample of 5k customers collected by our approach achieves lower profit loss than the one estimated from a sample of 10k customers collected by the two benchmarks. This efficiency is even more pronounced when the decision threshold further deviates from the mode (i.e., $c = 3$). These findings underscore the potential benefit of our proposed approach in significantly reducing experimentation costs without compromising targeting performance.

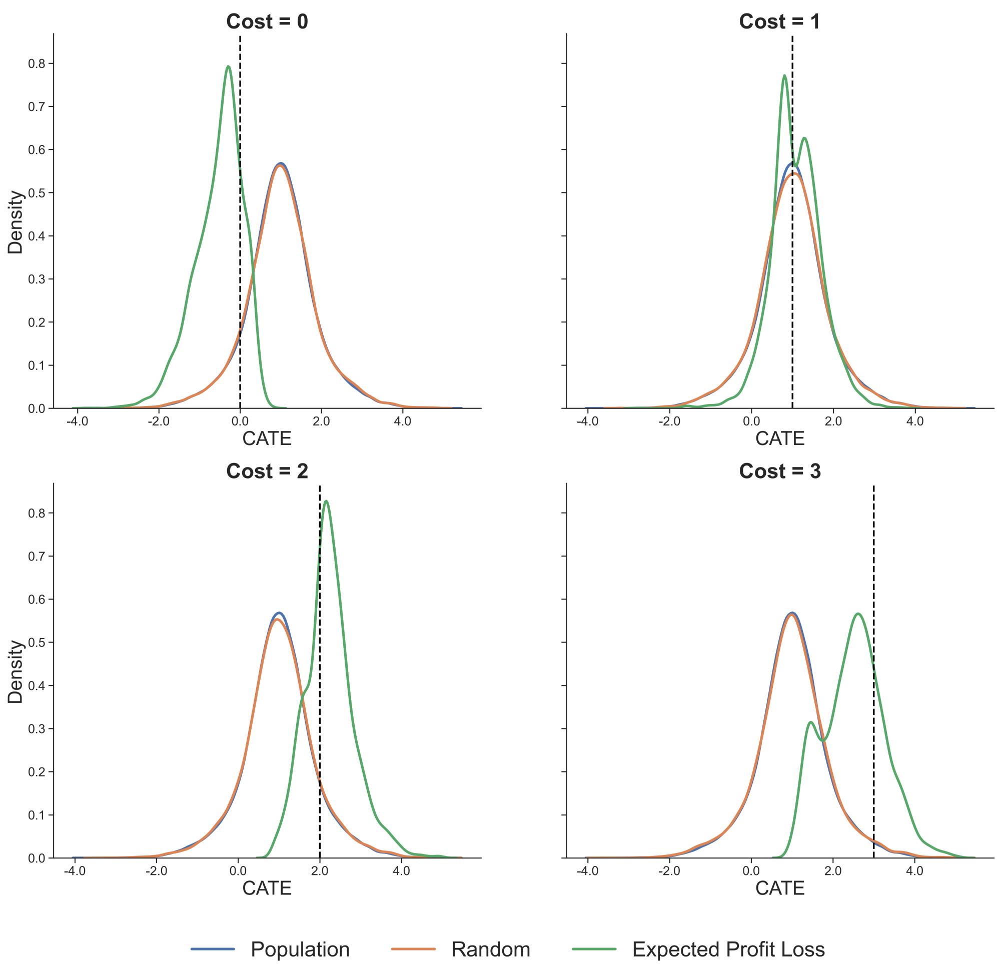
Finally, we underscore the effectiveness of our approach even in its simplified two-stage design. Specifically, the simplified two-stage version of our approach consistently outperforms the two benchmarks when $c \in \{0, 2, 3\}$, and performs comparably to the multiple-stage version. Moreover, when the cost is near the mode and the sample size is small (e.g., $N_e = 1000$), a two-stage design with a higher proportion of customers sampled in the first stage (i.e., $r = 0.9$) demonstrates the best performance among the four variants of our proposed approach. This is because in such scenarios,

[12] The results for $c = 0$ also suggest that our approach performs well in cases where the intervention is cost-free and customer responsiveness is high. However, in practice, a straightforward strategy of targeting all customers would already be near-optimal in such cases, leaving limited room for additional targeting improvements.

more accurate expected profit loss estimates are crucial to effectively sample the consequential customers as intensively as random sampling does. Hence, a greater proportion of customers sampled in the first stage becomes necessary when dealing with smaller sample sizes.

4.4.2. Effectiveness in Identifying Consequential Customers We further analyze the CATE distributions of the (sampled) customers to understand the effectiveness of each of the approaches in identifying the most consequential customers. Unlike real-world data, the simulation setting allows us to observe the true CATEs and therefore offers the opportunity to directly assess whether our sampling approaches successfully identify and select those customers whose CATE estimation errors are most consequential for targeting profitability.

Figure 7 CATE Distributions of Customers Sampled by Different Approaches



Note. Each line corresponds to the CATE distribution of the customers sampled by different approaches. The dashed line represents the intervention cost, which is also the decision threshold.

Figure 7 illustrates the CATE distributions of the customers sampled by different approaches (`Random` and `Expected profit loss`) together with the distribution of the entire population

(`Population`). Each subfigure presents one of the scenarios, depending on the intervention cost. [13] Consistently across all scenarios, the expected profit loss criterion selects customers whose CATEs are close to, yet slightly deviate from the decision threshold. Notably, when the cost aligns with the mode of the CATE distribution (i.e., $c = 1$), the CATE distribution of customers selected by our approach is even more concentrated around this mode than those selected by random sampling. Additionally, when the cost diverges from the mode (i.e., $c \in \{0, 2, 3\}$), our method results in a CATE distribution that concentrates around, though not directly at, the decision threshold.[14]

These results underscore the effectiveness of our approach in pinpointing consequential customers — those with CATEs near but not exactly at the decision threshold — and sampling them intensively. Conversely, customers selected through random sampling more closely resemble the population distribution. Consequently, the distinction in consequential customer sampling between our approach and random sampling becomes particularly pronounced when fewer customers are positioned near the decision threshold. This finding aligns with the results in Section 4.4, emphasizing the advantage of our method in boosting targeting performance under conditions with limited customers around the decision threshold.

In conclusion, our simulation results affirm the superiority of our approach in enhancing targeting performance and sample size efficiency, especially in scenarios where customers near the decision threshold are limited. This advantage arises from our method's ability to effectively identify consequential customers and prioritize their sampling, an outcome unattainable with random sampling alone. Furthermore, we show that even a simplified two-stage design of our approach maintains its efficacy, providing firms with a practical, efficient solution for improving targeting outcomes in cases where a fully adaptive design may be impractical.

5. Empirical Applications

5.1. Overview

In this section, we assess the performance of our proposed approach using two real-world applications: one application features a dormant reactivation campaign run by a telecommunication company and the other examines a promotional marketing campaign conducted by the coffee chain Starbucks. In both cases, we evaluate the targeting performance of two types of targeting policies,

[13] Here we set the sample size to 10k and implement the expected profit loss sampling strategy using a two-stage design with $r = 0.5$.

[14] When $c = 3$, our method tends to focus on sampling customers whose CATEs are just below the threshold, as the pool of customers with CATEs above the threshold is too limited for intensive sampling in this case. By contrast, when $c = 2$, there are a larger number of customers with CATEs above the threshold, enabling us to sample these customers more intensively as they are less likely to have been sampled in the first stage. Similarly, for $c = 0$, our approach intensifies sampling in regions where customers are less likely to have been selected initially.

the `Default` targeting policy, which is learned from randomly sampling customers for the experimental set and observing their outcomes, and the `Proposed` policy, which is learned by observing the outcomes of the customers selected via our proposed sampling strategy.[15] Consistent with the analyses in previous sections, we evaluate different variants of our approach by modifying the number of stages: from multiple batches (each with a size of 500) to a simplified two-stage design.[16] We also adjust the size of each batch in the two-stage condition. Finally, we test the performance of the different approaches when using alternative sample sizes of the experimental set, N_e, from 10,000 to 50,000.[17]

To assess the performance of each targeting approach, we use the bootstrap validation scheme similar to Ascarza (2018), where we generate 100 data splits, with each split consisting of a customer base, $\mathcal{I}$; 80%, and an evaluation set, D_{eval}; 20%. In each split, we sample N_e customers from the customer base, $\mathcal{I}$, following each of the sampling criteria (`Default` and `Proposed`). These sampled customers would constitute the experimental data, D_e, and will, therefore, be the data used to estimate the CATE model (for each approach). We then leverage the constructed CATE model to generate targeting decisions $\hat{\pi}(\cdot)$ on the evaluation set, D_{eval}.

We evaluate the targeting performance by first computing the expected profit generated by each estimated targeting policy, $\hat{\pi}(\cdot)$, as well as a uniform policy, π_u, that provides the treatment to every customer. Specifically, we use the inverse-probability-weighted (IPW) estimator (Horvitz and Thompson 1952, Hitsch and Misra 2018) to estimate the expected profit of a targeting policy, $\hat{\pi}(\cdot)$:

$$\text{Profit}(\hat{\pi}) = \sum_i \left(\frac{1-W_i}{1-e(\mathbf{X}_i)}(1-\hat{\pi}(\mathbf{X}_i))Y_i(0) + \frac{W_i}{e(\mathbf{X}_i)}\hat{\pi}(\mathbf{X}_i)(Y_i(1)-c_i) \right) \tag{6}$$

where $e(\mathbf{X}_i)$ is the (estimated) propensity score of customers assigned to the treatment condition in the evaluation set D_{eval}.

After obtaining the expected profit loss estimates, the proportional profit improvement of the targeting policy $\hat{\pi}(\cdot)$ relative to the uniform policy π_u (PIU) is then computed as the evaluation metric using the following formula:

$$\text{PPI} = \frac{\text{Profit}(\hat{\pi}) - \text{Profit}(\pi_u)}{\text{Profit}(\pi_u)} \tag{7}$$

[15] Given that the data has already been collected, we do not include Kato et al. (2024) as a benchmark in the empirical application. The primary reason is that the optimal treatment assignment rule from Kato et al. (2024) might allocate customers to different treatment conditions than those in the original experimental data. Consequently, implementing Kato et al. (2024) offline would necessitate accurately simulating customers' counterfactual behavior. Given the low response rates in both datasets, accurately predicting customers' counterfactual behavior in these contexts is very challenging.

[16] For the multi-stage design, we follow a similar procedure as Waisman et al. (2024) by using equal-sized batches throughout the experiment.

[17] The maximum value is determined by the total sample size of the data we collected from each case study.

where Profit(π_u) denotes the expected profit of the uniform policy. Since this metric quantifies the profit improvement of the estimated targeting policies relative to the uniform policy, a larger PPI indicates that the targeting decision estimated from the experimental data D_e is more profitable.

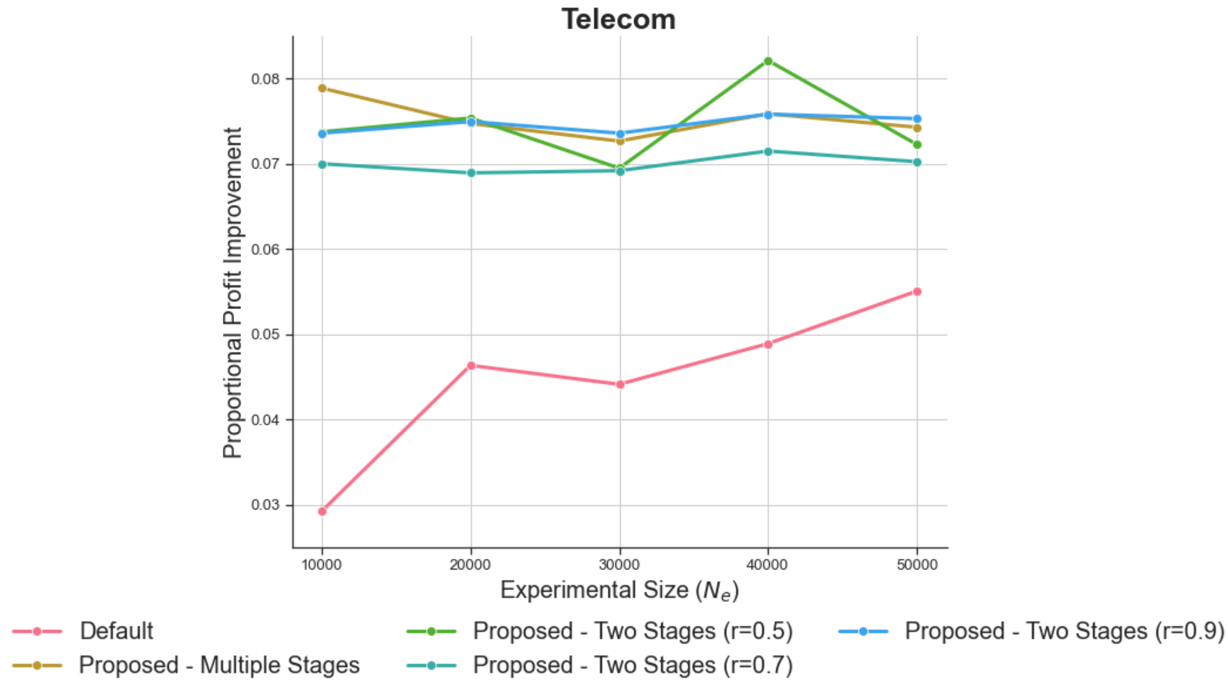
5.2. Telecommunication Dormant Reactivation Campaign

5.2.1. Empirical Context and Data Description Our first empirical application leverages the data from a dormant reactivation campaign conducted by a telecommunication company. The experimental data involves a reactivation campaign aiming to activate and increase users' usage over a 14-day period. The experiment ran from April 7^{th}, 2016 to January 10^{th}, 2018. Each week, the company identified eligible customers and randomly assigned them to either the control or treatment group. The campaign included 374,051 eligible customers, with 83,781 in the control group and 290,270 in the treatment group. Customers in the treatment group were offered 5 units of international voice credit for 3 days if they recharged at least 20 credit units within 3 days of receiving the offer. The outcome variable, a continuous measure, represented the total expenditure of each customer over the subsequent 14 days. While the intervention was cost-free for the company (i.e., giving free international voice credits doesn't cost the company anything), offering free international voice credits creates cannibalization as some of the customers would have otherwise paid for these credits. This cannibalization resulted in treated customers recharging less than those in the control group, leading to a negative average treatment effect of -0.65.

The dataset also includes various pre-treatment customer behaviors, which serve as covariates for targeting. These variables capture customers' previous behavior such as usage, recharge activity, and cancellation activity (of related services) over the past 7, 14, and 30 days. Additionally, because the focal company ran the campaign on a weekly basis, we created an additional variable, targeting ratio, which represents the proportion of customers targeted in a given week and is added as a control in all models. See Web Appendix D.1 for further details about the data.

5.2.2. Profitability of Targeting Policies We now turn to evaluate the performance of the different policies. Recall that for this analysis, we split the data into customer base, $\mathcal{I}$, used to run each experimental approach (default or proposed), and evaluation set, D_{eval}, used to evaluate the profitability of each policy, as defined in (7). Figure 8 shows the proportional profit improvement of the targeting policies relative to the uniform policy. Each line corresponds to the profitability of the policy derived by each experimental design (or sampling approach), across experimental sizes ($N_e \in \{10k, 20k, 30k, 40k, 50k\}$). In addition, we report the percentage of replications in which the proportional profit improvement of the focal method relative to the uniform policy is greater than the default approach in Table 1 to provide a better sense of the performance difference across the different approaches.

Figure 8 **Expected Profit of Targeting Policies Learned from Different Experimental Designs (Telecommunication)**



Note. We report the average value of the proportional profit improvement relative to the uniform policy across 100 replications. Each line corresponds to an experimental approach.

Table 1 **Percentage of Replications in which the Proposed Method Outperforms the Default Approach (Telecommunication)**

Experimental Design	**Experimental Size**				
	10k	20k	30k	40k	50k
Proposed Multiple Stages	74%	61%	69%	66%	59%
Proposed Two-Stage (r = 0.5)	75%	59%	63%	71%	61%
Proposed Two-Stage (r = 0.7)	69%	62%	68%	63%	59%
Proposed Two-Stage (r = 0.9)	70%	57%	64%	70%	59%

Note: We report the percentage of bootstrap replications in which the proportional profit improvement of the focal method relative to the uniform policy is greater than the default approach.

The results reveal that our approach significantly outperforms the default approach, especially when the experimental size is small. Remarkably, our method requires only 10,000 experimental samples to generate a more profitable targeting policy than the default method does with 50,000 samples. This demonstrates the practical value of our approach, especially in scenarios where increasing the experimental size is difficult or costly for firms, such as when the customer base is limited.

Additionally, the simplified two-stage designs achieve similar profitability to that of the fully adaptive design, regardless of the experimental size. Our findings highlight that even with a simplified implementation, the proposed method can enhance profitability, making it a valuable tool for firms with different capacities for experimental scale.

5.3. Starbucks Promotional Campaign

5.3.1. Empirical Context and Data Description Our second empirical application utilizes the data from a promotional campaign conducted through Starbucks' mobile reward app.[18] The experimental data involves a promotional campaign aiming to increase customers' purchase rate. In particular, the dataset contains 126,184 customers who were randomly assigned to either the control group (63,112) or the treatment group (63,072), with the treated customers receiving the promotional content offered by Starbucks. The outcome variable is binary, indicating whether the customer made a purchase or not. Notably, the response rates to the intervention are quite low, with 1.68% in the treatment group and 0.73% in the control group, leading to an average treatment effect of 0.95%. The data also includes seven pre-treatment covariates, which will be used for targeting.[19] See Web Appendix D.2 for further details about the data.

Because the data provided does not include specific details about the promotional content, we consider two scenarios. First, one in which Starbucks sent a promotional content without discount (e.g., a notification promoting Starbucks). This is the same as analyzing an intervention with, essentially, no cost or potential cannibalization. Second, we assume that the intervention sent to the treatment group was offering a 50% off discount in the next purchase. In this scenario, the intervention creates cannibalization since using the promotion implies lower revenue in the associated sale.[20] In both scenarios, we assume that each purchasing customer has an order value of $6, which is the average ticket size for a Starbucks customer.[21]. However, for customers who receive the 50% discount, their spending is reduced to $3 per order. Importantly, each scenario reflects different levels of cannibalization introduced by the intervention, enabling us to assess our approach's performance under varying relationships between the CATE distribution and the decision threshold, similar to our analysis in Section 4.

[18] The data is provided by Starbucks and was made available through the Udacity Data Science Program.

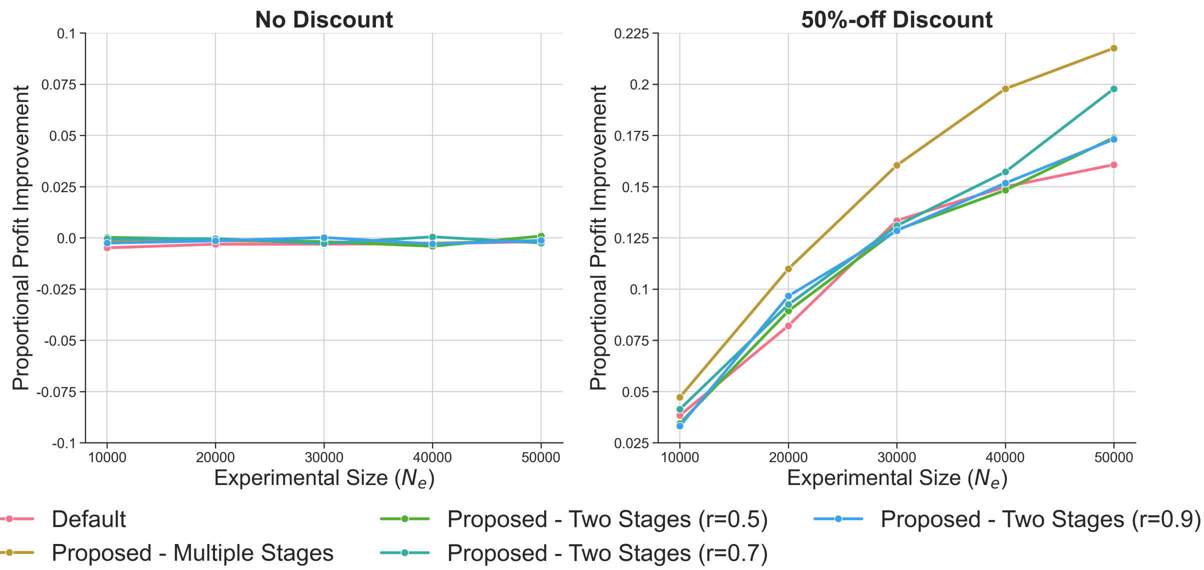
[19] The provided data is anonymized and does not include the meaning of each pre-treatment variable. While this limits the ability to interpret some findings, these variables contain the necessary information for determining targeting policies, which is our primary objective.

[20] We consider this case a cost-free intervention with a risk of cannibalization, rather than a costly intervention, due to the intervention's effect of reducing customer spending.

[21] Source: https://wifitalents.com/statistic/starbucks-customers/

5.3.2. Profitability of Targeting Policies We now turn to evaluate the performance of the different policies. Similar to the first empirical application, we split the data into customer base, $\mathcal{I}$, used to run each experimental approach (`Default` or `Proposed`), and evaluation set, D_{eval}, used to evaluate the profitability of each policy, as defined in (7). Figure 9 shows the proportional profit improvement of the targeting policies relative to the uniform policy for each scenario (no discount and 50%-off discount). Each line corresponds to the profitability of the policy derived by each experimental design (or sampling approach), across experimental sizes ($N_e \in$ {10k, 20k, 30k, 40k, 50k}). In addition, for the scenario where Starbucks offer a 50%-off discount, we report the percentage of replications in which the proportional profit improvement of the focal method relative to the uniform policy is greater than the default approach in Table 2 to provide a better sense of the performance difference across different approaches.

Figure 9 Expected Profit of Targeting Policies Learned from Different Experimental Designs (Starbucks)



Note. We report the average value of the proportional profit improvement relative to the uniform policy across 100 replications. Each line corresponds to an experimental approach.

We highlight several interesting findings. First, our approach significantly outperforms the default approach when Starbucks offers a 50%-off discount (i.e., when the number of customers around the decision threshold is limited) while it performs at par with the benchmark approach in the "No Discount" case. This finding is consistent with our simulation results and underscores the practical benefit of our method in enhancing the profitability of targeting policies, especially in scenarios

Table 2 Percentage of Replications in which the Proposed Method Outperforms the Default Approach (Starbucks 50%-off Discount)

Experimental Design	Experimental Size				
	10k	20k	30k	40k	50k
Proposed Multiple Stages	51%	60%	63%	74%	72%
Proposed Two-Stage (r = 0.5)	49%	58%	47%	52%	60%
Proposed Two-Stage (r = 0.7)	53%	55%	45%	56%	81%
Proposed Two-Stage (r = 0.9)	52%	57%	48%	49%	60%

Note: We report the percentage of bootstrap replications in which the proportional profit improvement of the focal method relative to the uniform policy is greater than the default approach when Starbucks offers a 50%*-off discount.*

where the intervention cost is high, but customer responsiveness might be low, as is the case when Starbucks offers a 50%-off discount.[22]

Second, even when the proposed approach does not provide additional benefits in learning targeting policies, as in Scenario 1, our method (whether fully adaptive or performed in two stages) does not harm profitability. This highlights the low risk of implementing the proposed approach in practice.

Third, when comparing the performance of targeting policies across different experimental sizes, the default method requires at least $50k$ experimental samples to achieve the same level of profitability as the fully adaptive method with only $30k$ experimental samples. This underscores the value of the proposed approach, particularly in situations where increasing the experimental size is challenging or costly for firms, such as when the customer base size is limited.

5.4. Discussion

Across both empirical applications, our proposed method consistently shows either comparable or superior performance relative to the default approach. Notably, it excels when the intervention creates a risk of cannibalizing profits that would have been earned in the absence of intervention, as evidenced in both the telecommunications application (offering free credits) and the 50% discount scenario for Starbucks (offering discounts). In these cases, the interventions reduce incremental

[22] In the "No Discount" scenario, the estimated targeting policies perform similarly to the uniform policy across different experimental designs. This suggests that when the intervention is cost-free and customers respond positively, a simple strategy of targeting all customers would already be close to an optimal approach.

spending for targeted customers, resulting in a negative average treatment effect. This result arises because, with a combination of weak treatment effects and cannibalization caused by the intervention, the CATE distribution tends to be leftward shifted (negative treatment effect). However, the firm's decision threshold still falls at zero. This misalignment between the mode of the CATE distribution and the decision threshold results in fewer customers positioned near the threshold, making it harder to identify critical customers effectively.

Notably, this scenario frequently occurs in various marketing campaigns involving free goods or monetary incentives, such as coupons (e.g., Ascarza 2018, Yang et al. 2023). In addition, as predicted by our simulation results, our approach is also expected to be beneficial in cases where there is a misalignment between intervention costs and the central tendency of customer responsiveness, such as when (1) the intervention is detrimental for most customers (e.g., Ascarza et al. 2016), or (2) the intervention incurs costs (e.g., phone calls, mailings) but shows low effectiveness (e.g., Lemmens and Gupta 2020). In general, we recommend firms adopt our approach whenever there is a likely disparity between intervention cost and customer responsiveness.[23]

6. Conclusion

With unprecedented access to consumer information, firms are increasingly interested in designing highly effective data-driven targeting policies based on detailed consumer data. The current "test-then-learn" approach provides firms with a method to leverage randomized experiments to predict consumer responsiveness to marketing interventions and design targeted policies accordingly. However, these experiments are often designed to obtain representative samples to assess the intervention's treatment effect, which may not align with the firm's primary business objective (e.g., maximizing targeting profitability). Such misalignment between experimental design and firms' business objectives can impede their ability to optimize business outcomes, ultimately reducing the effectiveness of their targeting strategies.

To address this issue, we propose a sequential experimental design coupled with a novel sampling criterion — expected profit loss — that integrates firms' profit-maximizing objectives directly into the experimental design. Specifically, rather than randomly sampling all customers at once, we propose dividing the predetermined experimental size into several smaller batches. Customers are then sampled sequentially based on the highest expected profit loss estimated from previous batches, allowing the firm to oversample those whose prediction errors in CATE estimation most significantly impact profitability. We demonstrate theoretically that our expected profit loss sampling

[23] Theoretically, our approach is also expected to perform well when the intervention is costless and customers are highly responsive, as the decision threshold in such cases similarly deviates from the mode of the CATE distribution. However, in these scenarios, a straightforward strategy that targets every customer would already be close to optimal, offering limited room for further improvement in targeting performance as shown in Section 5.3.2

strategy aligns with the firms' profit-maximizing objectives and introduce a novel expected profit loss estimation method based on Causal Forest (Wager and Athey 2018, Athey et al. 2019b). Our solution allows firms to enhance targeting policy effectiveness without increasing the experimental size by focusing on consequential customers who significantly impact profitability.

Through both simulation studies and empirical applications, we demonstrate the superiority of our proposed approach in identifying the most profitable targeting policies compared to the current test-then-learn practice and the state-of-the-art adaptive experimental design, especially when there are few customers around the decision threshold. This typically occurs when (1) intervention is harmful for most customers (e.g. Ascarza et al. 2016), (2) the intervention is costly but ineffective (e.g. Lemmens and Gupta 2020, Simester et al. 2022), or (3) there is a risk of cannibalizing revenues/profits (e.g. Anderson and Simester 2004, Ascarza 2018, Yang et al. 2023), as in our empirical applications.

Our results also highlight the efficacy of our approach when simplified into a two-stage design. This streamlined approach divides the experimental sample into just two groups, accelerating the experimentation process by requiring only two rounds of customer response collection. This simplification is particularly beneficial when there is a delay between interventions and customer responses. Moreover, the two-stage design minimizes the need for repeated estimation and adjustment of the expected profit loss, thereby lowering implementation costs and making the approach more appealing to marketing practitioners.

Although our research provides a simple and efficient solution for firms to improve targeted policies, there are limitations that suggest promising directions for future research. First, we focus on scenarios where the firms' profit-maximizing objectives are not subject to any constraints. However, in practice, some firms may face several managerial constraints on their targeted policies, such as budget constraints or fairness constraints (Lu et al. 2023). Future research could build on the decision-aware learning literature (e.g. Chung et al. 2022, Liu et al. 2023) to explore different sampling strategies that integrate firms' business objectives into the experimental design while accounting for these constraints.

Second, our approach assumes that firms lack prior knowledge about consumers' responsiveness to interventions. However, firms often possess historical experimental or observational data that, while related, may differ from the current experiment. For instance, firms may have conducted campaigns involving different interventions, such as varying discount levels. While previous research has explored transferring knowledge from past marketing campaigns to enhance targeting policies (Timoshenko et al. 2020, Huang et al. 2024) it has not addressed the misalignment between experimentation approaches and the firm's objectives. Future research could investigate ways to incorporate information from previous campaigns to refine the design of focal experiments in a policy-aware manner.

Third, although our two-stage design remains effective when there is a delay between the intervention and the outcome of interest, more efficient methods could address delayed feedback. Prior work has examined strategies for handling delayed feedback in multi-armed bandits by explicitly modeling the relationship between partial and delayed feedback (e.g. Grover et al. 2018). Building on these approaches, future research could extend our proposed sequential design to better manage significantly delayed feedback and reduce the overall lag time.

Fourth, our approach assumes that firms have a fixed customer base, allowing them to easily determine which customers warrant closer attention based on the potential impact of CATE estimation errors on profitability. While this scenario is common in various marketing applications such as customer retention (e.g., Ascarza 2018, Lemmens and Gupta 2020, Yang et al. 2023) and most promotional activities to incentivize consumption (e.g., Hitsch and Misra 2018, Simester et al. 2020), there are instances where customers arrive sequentially and unpredictably, such as in digital advertising or customer acquisition strategies. In these cases, strategic sampling becomes more challenging, as firms cannot anticipate whether future customers might be more consequential and thus deserve greater focus. Future research could draw upon insights from the online active learning literature (Cacciarelli and Kulahci 2024) to investigate optimal sampling strategies in such dynamic environments.

Finally, alternative enhancements to our policy-aware approach could be explored. For instance, we proposed selectively sampling customers while assigning treatments randomly. Future research could investigate methods for both selectively sampling customers and strategically assigning them to treatment conditions to better align with firms' business objectives and improve targeting performance. Additionally, in cases where control condition is equivalent to business as usual, one could extend our work to use all non-experimental customers as a control group to enhance the estimation efficiency. Future research could explore the optimal sampling strategy under this framework.

Overall, our research demonstrates the value of incorporating firms' business objective into the design of experiments. We hope that our work will inspire further research on aligning the science of experimentation with firms' objectives across various scenarios.

References

Ailawadi KL, Harlam BA, César J, Trounce D (2007) Practice Prize Report—Quantifying and Improving Promotion Effectiveness at CVS. *Marketing Science* 26(4):566–575.

Anderson ET, Simester DI (2004) Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies. *Marketing Science* 23(1):4–20.

Aramayo N, Schiappacasse M, Goic M (2022) A Multiarmed Bandit Approach for House Ads Recommendations. *Marketing Science* mksc.2022.1378.

Ascarza E (2018) Retention Futility: Targeting High-Risk Customers Might be Ineffective. *Journal of Marketing Research* 55(1):80–98.

Ascarza E, Iyengar R, Schleicher M (2016) The Perils of Proactive Churn Prevention Using Plan Recommendations: Evidence from a Field Experiment. *Journal of Marketing Research* 53(1):46–60.

Athey S, Byambadalai U, Hadad V, Krishnamurthy SK, Leung W, Williams JJ (2022) Contextual Bandits in a Survey Experiment on Charitable Giving: Within-Experiment Outcomes versus Policy Learning. *arXiv preprint arXiv:2211.12004* .

Athey S, Chetty R, Imbens GW, Kang H (2019a) The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. Working Paper 26463, National Bureau of Economic Research.

Athey S, Tibshirani J, Wager S (2019b) Generalized random forests. *The Annals of Statistics* 47(2).

Athey S, Wager S (2021) Policy Learning With Observational Data. *Econometrica* 89(1):133–161.

Blattberg RC (1979) The Design of Advertising Experiments Using Statistical Decision Theory. *Journal of Marketing Research* 16(2):191.

Bubeck S, Munos R, Stoltz G (2010) Pure Exploration for Multi-Armed Bandit Problems. *arXiv preprint arXiv:0802.2655* .

Cacciarelli D, Kulahci M (2024) Active learning for data streams: a survey. *Machine Learning* 113(1):185–239.

Cardoso TN, Silva RM, Canuto S, Moro MM, Gonçalves MA (2017) Ranked batch-mode active learning. *Information Sciences* 379:313–337.

Caria S, Gordon G, Kasy M, Quinn S, Shami S, Teytelboym A (2020) An Adaptive Targeted Field Experiment: Job Search Assistance for Refugees in Jordan. *Available at SSRN 3689456* .

Carranza AG, Krishnamurthy SK, Athey S (2023) Flexible and Efficient Contextual Bandits with Heterogeneous Treatment Effect Oracles. *arXiv preprint arXiv:2203.16668* .

Chick SE, Frazier P (2012) Sequential Sampling with Economics of Selection Procedures. *Management Science* 58(3):550–569.

Chung TH, Rostami V, Bastani H, Bastani O (2022) Decision-Aware Learning for Optimizing Health Supply Chains. *arXiv preprint arXiv:2211.08507* .

Dew R (2023) Adaptive Preference Measurement with Unstructured Data. *Available at SSRN 4641773* .

Dzyabura D, Hauser JR (2011) Active Machine Learning for Consideration Heuristics. *Marketing Science* 30(5):801–819.

Ellickson PB, Kar W, Reeder JC (2022) Estimating Marketing Component Effects: Double Machine Learning from Targeted Digital Promotions. *Marketing Science* mksc.2022.1401.

Elmachtoub AN, Grigas P (2022) Smart "Predict, then Optimize". *Management Science* 68(1):9–26.

Feit EM, Berman R (2019) Test & Roll: Profit-Maximizing A/B Tests. *Marketing Science* 38(6):1038–1058.

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

Filstroff L, Sundin I, Mikkola P, Tiulpin A, Kylmäoja J, Kaski S (2021) Targeted Active Learning for Bayesian Decision-Making. *arXiv preprint arXiv:2106.04193* .

Fu Y, Zhu X, Li B (2013) A survey on instance selection for active learning. *Knowledge and Information Systems* 35(2):249–283.

Ginter JL, Cooper MC, Obermiller C, Page TJ (1981) The Design of Advertising Experiments Using Statistical Decision Theory: An Extension. *Journal of Marketing Research* 18(1):120–123.

Grover A, Markov T, Attia P, Jin N, Perkins N, Cheong B, Chen M, Yang Z, Harris S, Chueh W, Ermon S (2018) Best arm identification in multi-armed bandits with delayed feedback. *arXiv preprint arXiv:1803.10937* .

Hadad V, Rosenzweig LR, Athey S, Karlan D (2021) Designing Adaptive Experiments. Practitioner's Guide, Stanford Business School.

Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning*. Springer Series in Statistics (New York, NY: Springer New York).

Hauser JR, Urban GL, Liberali G, Braun M (2009) Website Morphing. *Marketing Science* 28(2):202–223.

Hitsch GJ, Misra S (2018) Heterogeneous Treatment Effects and Optimal Targeting Policy Evaluation. *Available at SSRN 3111957* .

Holland PW (1986) Statistics and causal inference. *Journal of the American Statistical Association* 81(396):945–960.

Horvitz DG, Thompson DJ (1952) A Generalization of Sampling Without Replacement from a Finite Universe. *Journal of the American Statistical Association* 47(260):663–685.

Hu Y, Zhu H, Brunskill E, Wager S (2024) Minimax-Regret Sample Selection in Randomized Experiments. *arXiv preprint arXiv:2403.01386* .

Huang TW, Ascarza E (2024) Doing More with Less: Overcoming Ineffective Long-Term Targeting Using Short-Term Signals. *Marketing Science* mksc.2022.0379.

Huang TW, Ascarza E, Israeli A (2024) Incrementality representation learning: Synergizing past experiments for intervention personalization. *Available at SSRN 4859809* .

Jain L, Li Z, Loghmani E, Mason B, Yoganarasimhan H (2024) Effective Adaptive Exploration of Prices and Promotions in Choice-Based Demand Models. *Marketing Science* mksc.2023.0322.

Jedra Y, Proutiere A (2020) Optimal Best-arm Identification in Linear Bandits. *arXiv preprint arXiv:2006.16073* .

Jesson A, Tigas P, van Amersfoort J, Kirsch A, Shalit U, Gal Y (2022) Causal-BALD: Deep Bayesian Active Learning of Outcomes to Infer Treatment-Effects from Observational Data. *arXiv preprint arXiv:2111.02275* .

Kallus N, Mao X (2023) Stochastic Optimization Forests. *Management Science* 69(4):1975–1994.

Kasy M, Sautmann A (2021) Adaptive Treatment Assignment in Experiments for Policy Choice. *Econometrica* 89(1):113–132.

Kato M, Okumura K, Ishihara T, Kitagawa T (2024) Adaptive Experimental Design for Policy Learning. *arXiv preprint arXiv:2401.03756* .

Kotary J, Fioretto F, Van Hentenryck P, Wilder B (2021) End-to-End Constrained Optimization Learning: A Survey. *arXiv preprint arXiv:2103.16378* .

Lemmens A, Gupta S (2020) Managing Churn to Maximize Profits. *Marketing Science* 39(5):956–973.

Li L, Chu W, Langford J, Schapire RE (2010) A Contextual-Bandit Approach to Personalized News Article Recommendation. *Proceedings of the 19th international conference on World wide web*, 661–670.

Liu M, Grigas P, Liu H, Shen ZJM (2023) Active Learning in the Predict-then-Optimize Framework: A Margin-Based Approach. *arXiv preprint arXiv:2305.06584* .

Lu H, Simester D, Zhu Y (2023) Optimizing Scalable Targeted Marketing Policies with Constraints. *Available at SSRN 4668582* .

Misra K, Schwartz EM, Abernethy J (2019) Dynamic Online Pricing with Incomplete Information Using Multiarmed Bandit Experiments. *Marketing Science* 38(2):226–252.

Nguyen X, Wainwright MJ, Jordan MI (2009) On surrogate loss functions and f-divergences. *The Annals of Statistics* 37(2):876–904.

Puha Z, Kaptein M, Lemmens A (2020) Batch Mode Active Learning for Individual Treatment Effect Estimation. *2020 International Conference on Data Mining Workshops (ICDMW)*, 859–866 (Sorrento, Italy: IEEE), ISBN 978-1-72819-012-9.

Robinson PM (1988) Root-N-Consistent Semiparametric Regression. *Econometrica* 56(4):931.

Schwartz EM, Bradlow ET, Fader PS (2017) Customer Acquisition via Display Advertising Using Multi-Armed Bandit Experiments. *Marketing Science* 36(4):500–522.

Simester D, Timoshenko A, Zoumpoulis S (2022) A Sample Size Calculation for Training and Certifying Targeting Policies. *Available at SSRN 4228297* .

Simester D, Timoshenko A, Zoumpoulis SI (2020) Targeting Prospective Customers: Robustness of Machine-Learning Methods to Typical Data Challenges. *Management Science* 66(6):2495–2522.

Sundin I, Schulam P, Siivola E, Vehtari A, Saria S, Kaski S (2019) Active Learning for Decision-Making from Imbalanced Observational Data. *arXiv preprint arXiv:1904.05268* .

Timoshenko A, Ibragimov M, Simester D, Parker J, Schoar A (2020) Transferring information between marketing campaigns to improve targeting policies. Technical report, Working Paper.

Wager S, 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 HS, Carrion C (2024) Online Causal Inference for Advertising in Real-Time Bidding Auctions. *Marketing Science* mksc.2022.0406.

Wang K, Wilder B, Perrault A, Tambe M (2020) Automatically learning compact quality-aware surrogates for optimization problems. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, 9586–9596.

Wang Z, Ye J (2015) Querying Discriminative and Representative Samples for Batch Mode Active Learning. *ACM Transactions on Knowledge Discovery from Data* 9(3):1–23.

Wilder B, Dilkina B, Tambe M (2019) Melding the Data-Decisions Pipeline: Decision-Focused Learning for Combinatorial Optimization. *Proceedings of the AAAI Conference on Artificial Intelligence* 33(01):1658–1665.

Yang J, Eckles D, Dhillon P, Aral S (2023) Targeting for Long-Term Outcomes. *Management Science* mnsc.2023.4881.

Yoganarasimhan H, Barzegary E, Pani A (2023) Design and Evaluation of Optimal Free Trials. *Management Science* 69(6):3220–3240.

Zhang T (2004) Statistical behavior and consistency of classification methods based on convex risk minimization. *The Annals of Statistics* 32(1):56–85.

Web Appendix A: Comparison of Methodologies in Literature

Table App-1 Comparison of Methodologies in Literature

Literature	Objective	Sampling Strategy	Sampling Space	Pre-treatment Covariates	Pre-Segmentation	Sequential	Examples
Feit and Berman (2019)	Calculate sample size	N/A	N/A	N	N	N	N/A
Simester et al. (2022)	Calculate sample size	N/A	N/A	Y	Y	N	N/A
Hu et al. (2024)	Learning minimax-regret policy	Stratified Sampling	Context	Y	Y	N	N/A
Multi-Armed Bandit	Balancing exploration and exploitation	TS, UCB, etc.	Action	N	N	Y	Schwartz et al. (2017), Misra et al. (2019), **?**
Contextual Bandit	Balancing exploration and exploitation	TS, UCB, etc.	Action	Y	N	Y	Hauser et al. (2009), Li et al. (2010), Caria et al. (2020), Aramayo et al. (2022), Athey et al. (2022)
Best Arm Identification (BAI)	Identify best treatment with minimum sample	TS, UCB, etc.	Action	N	N	Y	Bubeck et al. (2010), Chick and Frazier (2012), Grover et al. (2018), Kasy and Sautmann (2021)
Contextual BAI	Learning personalized policy with minimum sample	TS, UCB, etc.	Action	Y	N	Y	Jedra and Proutiere (2020), Carranza et al. (2023), Kato et al. (2024)
Waisman et al. (2024)	Reducing experimentation Cost for ad effect estimation	TS	Action	Y	N	Y	N/A
Preference Measurement	Learning consumer preference with minimum questions	Bayesian Optimization	Action	N	N	Y	Dzyabura and Hauser (2011), Dew (2023)
Active Learning for Supervised Learning	Optimize data collection for supervised learning	Uncertainty Sampling, BALD, etc.	Context	Y	N	Y	Fu et al. (2013), Wang and Ye (2015), Cardoso et al. (2017)
Active Learning for CATE Estimation	Optimize data collection for CATE estimation	EMCM, BALD	Context	Y	N	Y	Puha et al. (2020), Jesson et al. (2022)
Active Learning for Decision Making	Optimize data collection for decision making	Uncertainty Sampling	Context	Y	N	Y	Sundin et al. (2019), Filstroff et al. (2021)
Active Learning for Predict-then-Optimize	Optimize data collection for predict-then-optimize	Margin-based Sampling	Context	Y	N	Y	Liu et al. (2023)
Our approach	Learning profit-maximizing targeting policy with pre-determined sample size	Expected Profit Loss Sampling	Context	Y	N	Y	N/A

Web Appendix B: Proof of Proposition 1

To prove Proposition 1, we first prove the following lemma:

Lemma 1 **(Consistency of logistic loss function)**. *The CATE estimate* $\hat{\tau}(\mathbf{X}_i)$ *that minimizes the logistic loss function*

$$\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i)) = \mathbb{E}[|\tau(\mathbf{X}_i)-c(\mathbf{X}_i)|\cdot\log_2(1+\exp(-\tilde{\pi}(\mathbf{X}_i)\cdot(\hat{\tau}(\mathbf{X}_i)-c(\mathbf{X}_i))))] \tag{App-1}$$

also minimizes the following objective function

$$\mathcal{L}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i)) = \mathbb{E}\left[|\tau(\mathbf{X}_i)-c(\mathbf{X}_i)|\cdot\mathbf{1}\{\pi^*(\mathbf{X}_i)\neq\hat{\pi}(\mathbf{X}_i)\}\right] \tag{App-2}$$

where

$$\pi^*(\mathbf{X}_i) = \mathbf{1}\{\tau(\mathbf{X}_i) > c(\mathbf{X}_i)\}$$

$$\hat{\pi}(\mathbf{X}_i) = \mathbf{1}\{\hat{\tau}(\mathbf{X}_i) > c(\mathbf{X}_i)\}$$

$$\tilde{\pi}(\mathbf{X}_i) = \begin{cases} 1 & \text{if } \pi^*(\mathbf{X}_i)=1 \\ -1 & \text{if } \pi^*(\mathbf{X}_i)=0 \end{cases}$$

Proof: Since the CATE estimate $\hat{\tau}(\mathbf{X}_i)$ that minimizes (App-2) satisfies

$$\mathbf{1}\{\pi^*(\mathbf{X}_i)\neq\hat{\pi}(\mathbf{X}_i)\} = 0$$

implying that

$$(\tau(\mathbf{X}_i)-c(\mathbf{X}_i))\cdot(\hat{\tau}(\mathbf{X}_i)-c(\mathbf{X}_i)) > 0.$$

It suffices to show that

$$(\tau(\mathbf{X}_i)-c(\mathbf{X}_i))\cdot(\hat{\tau}^*(\mathbf{X}_i)-c(\mathbf{X}_i)) > 0$$

where

$$\hat{\tau}^*(\mathbf{X}_i) = \arg\min_{\hat{\tau}}\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i));\tau(\mathbf{X}_i).$$

Observe that the CATE estimate $\hat{\tau}^*(\mathbf{X}_i)$ that minimizes $\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$ satisfies

$$-\tilde{\pi}(\mathbf{X}_i)\cdot(\hat{\tau}^*(\mathbf{X}_i)-c(\mathbf{X}_i)) \geq 0,$$

we have

$$\tilde{\pi}(\mathbf{X}_i)\cdot(\hat{\tau}^*(\mathbf{X}_i)-c(\mathbf{X}_i)) > 0.$$

By definition of $\tilde{\pi}(\mathbf{X}_i)$, this implies that

$$(\tau(\mathbf{X}_i)-c(\mathbf{X}_i))\cdot(\hat{\tau}(\mathbf{X}_i)-c(\mathbf{X}_i)) > 0.$$

This completes the proof.

We now prove proposition 1:

Proof of Proposition 1: Given $\mathbf{X}_i$, the firm's objective is to maximize the expected profit of $\hat{\pi}(\mathbf{X}_i)$, which is equivalent to minimizing the expected profit loss of $\hat{\pi}(\mathbf{X}_i)$ relative to the true optimal policy $\pi^*(\mathbf{X}_i)$:

$$\begin{aligned}\mathcal{L}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i)) &= \mathbb{E}\left[\mathbb{E}[\underbrace{(Y_i(0)+(Y_i(1)-Y_i(0)-c_i)\cdot\pi^*(\mathbf{X}_i))}_{\text{profit of }\pi^*(\mathbf{X}_i)} - \underbrace{(Y_i(0)+(Y_i(1)-Y_i(0)-c_i)\cdot\hat{\pi}(\mathbf{X}_i)}_{\text{profit of }\hat{\pi}(\mathbf{X}_i)}|\mathbf{X}_i]\right] \\ &= \mathbb{E}\left[|\tau(\mathbf{X}_i)-c(\mathbf{X}_i)|\cdot\mathbf{1}\{\pi^*(\mathbf{X}_i)\neq\hat{\pi}(\mathbf{X}_i)\}\right]\end{aligned}$$

where

$$\pi^*(\mathbf{X}_i) = \mathbf{1}\{\tau(\mathbf{X}_i) > c(\mathbf{X}_i)\}$$

$$\hat{\pi}(\mathbf{X}_i) = \mathbf{1}\{\hat{\tau}(\mathbf{X}_i) > c(\mathbf{X}_i)\}.$$

Our goal is to find the CATE estimates $\hat{\tau}(\cdot)$ such that the policy $\hat{\pi}(\cdot)$ generated from them minimizes the objective function $\mathcal{L}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$. However, since $\mathcal{L}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$ is discontinuous and non-convex and thus difficult to minimize, we leverage logistic loss $\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$ as a surrogate loss function for $\mathcal{L}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$ (Zhang 2004, Nguyen et al. 2009), i.e.,

$$\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i)) = \mathbb{E}[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \log_2(1 + \exp(-\tilde{\pi}(\mathbf{X}_i) \cdot (\hat{\tau}(\mathbf{X}_i) - c(\mathbf{X}_i))))]$$

where

$$\tilde{\pi}(\mathbf{X}_i) = \begin{cases} 1 \text{ if } \pi^*(\mathbf{X}_i) = 1 \\ -1 \text{ if } \pi^*(\mathbf{X}_i) = 0 \end{cases}$$

Since $\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$ is differentiable, by Taylor approximation, we have

$$\begin{aligned}
\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i)) &\approx \underbrace{\tilde{\mathcal{L}}(\tau(\mathbf{X}_i);\tau(\mathbf{X}_i))}_{\text{constant}} + \mathbb{E}\left[\nabla_{\hat{\tau}}\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))|_{\hat{\tau}=\tau} \cdot (\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i))\right] \\
&= \text{constant} + \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \frac{\exp(-\tilde{\pi}(\mathbf{X}_i)\cdot(\tau(\mathbf{X}_i) - c(\mathbf{X}_i)))}{1+\exp(-\tilde{\pi}(\mathbf{X}_i)\cdot(\tau(\mathbf{X}_i) - c(\mathbf{X}_i)))} \cdot (-\tilde{\pi}(\mathbf{X}_i)) \cdot (\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i))\right] \\
&= \text{constant} + \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \frac{1}{1+\exp(\tilde{\pi}(\mathbf{X}_i)\cdot(\tau(\mathbf{X}_i) - c(\mathbf{X}_i)))} \cdot (-\tilde{\pi}(\mathbf{X}_i)) \cdot (\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i))\right] \\
&= \text{constant} + \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \frac{1}{1+\exp(|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|)} \cdot (-\tilde{\pi}(\mathbf{X}_i)) \cdot (\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i))\right] \\
&\leq \text{constant} + \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \frac{1}{1+\exp(|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|)} \cdot |\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)|\right]
\end{aligned} \tag{App-3}$$

since if $\tilde{\pi}(\mathbf{X}_i) = 1$, $\tau(\mathbf{X}_i) - c(\mathbf{X}_i) > 0$; if $\tilde{\pi}(\mathbf{X}_i) = -1$, $\tau(\mathbf{X}_i) - c(\mathbf{X}_i) \leq 0$.

(App-3) implies that to minimize $\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i);\tau(\mathbf{X}_i))$, we would want to minimize the absolute prediction error $|\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)|$ more for the customers with larger

$$|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \frac{1}{1+\exp(|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|)}. \tag{App-4}$$

We now show that customers with larger values of (App-4) are identical to those with higher expected profit loss. By Huang and Ascarza (2024), we know that if $\tau(\mathbf{X}_i) > c(\mathbf{X}_i)$:

$$\begin{aligned}
P\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\} &= P\{\hat{\tau}(\mathbf{X}_i) - c(\mathbf{X}_i) < 0\} \\
&= P\left\{\frac{\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x} < \frac{c(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x}\right\} \\
&= \mathcal{F}_Z\left(\frac{c(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x}\right)
\end{aligned}$$

where

$$Z = \frac{\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x}$$

is a random variable with mean 0 and variance 1. Similarly, if $\tau(\mathbf{X}_i) < c(\mathbf{X}_i)$:

$$\begin{aligned}
P\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\} &= P\{\hat{\tau}(\mathbf{X}_i) - c(\mathbf{X}_i) > 0\} \\
&= P\left\{\frac{\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x} > \frac{c(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x}\right\} \\
&= 1 - \mathcal{F}_Z\left(\frac{c(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x}\right)
\end{aligned}$$

where

$$Z = \frac{\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)}{\sigma_x}$$

is a random variable with mean 0 and variance 1. Therefore, we can see that the probability of mistargeting $P\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\}$ decreases with $|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|$, implying that the probability of mistargeting $P\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\}$ increases with

$$\frac{1}{1 + \exp(|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|)}.$$

Thus, we can conclude that the customers with larger

$$|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \frac{1}{1 + \exp(|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)|)}.$$

are those with higher expected profit loss

$$\begin{aligned} |\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot P\{\hat{\pi}(\mathbf{X}_i) \neq \pi^*(\mathbf{X}_i)\} &= |\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \mathbb{E}\left[\mathbf{1}\{\pi^*(\mathbf{X}_i) \neq \hat{\pi}(\mathbf{X}_i)\}\right] \\ &= \mathbb{E}\left[|\tau(\mathbf{X}_i) - c(\mathbf{X}_i)| \cdot \mathbf{1}\{\pi^*(\mathbf{X}_i) \neq \hat{\pi}(\mathbf{X}_i)\}\right] \end{aligned}$$

, implying that to minimize $\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i); \tau(\mathbf{X}_i))$, we should focus on minimizing the absolute prediction error $|\hat{\tau}(\mathbf{X}_i) - \tau(\mathbf{X}_i)|$ for the customers with higher expected profit loss.

Since by Lemma 1, the CATE estimates $\hat{\tau}(\mathbf{X}_i)$ that minimizes $\tilde{\mathcal{L}}(\hat{\tau}(\mathbf{X}_i); \tau(\mathbf{X}_i))$ also minimizes the original objective function $\mathcal{L}(\hat{\tau}(\mathbf{X}_i); \tau(\mathbf{X}_i))$, we can conclude that to minimize the objective function $\mathcal{L}(\hat{\tau}(\mathbf{X}_i); \tau(\mathbf{X}_i))$ (i.e., to maximize the profit), the firm should focus on minimizing the absolute prediction error of the CATE estimates $\hat{\tau}(\mathbf{X}_i)$ for the customers with higher expected profit loss.

Web Appendix C: Further Details of Simulation Studies

C.1. Implementation Details

C.1.1. Default (A/B Test with Random Sampling) For each experimental size $N_e \in \{1k, 5k, 10k, 20k, 30k\}$, we randomly sample N_e customers from the customer base $\mathcal{I}$ and assign them randomly to two treatment conditions with a probability of 0.5. For CATE estimation, we construct a Causal Forest model (Wager and Athey 2018) implemented using the `econML` package in Python. This model consists of 100 trees with a maximum depth of 10 to prevent overfitting.

C.1.2. Kato et al. (2024) For each experimental size $N_e \in \{1k, 5k, 10k, 20k, 30k\}$, we follow a similar procedure as Waisman et al. (2024) by using equal-sized batches throughout the experiment. Specifically, we partition the full sample into $\frac{N_e}{200}$ batches, and randomly sample 200 customers who have not been sampled in previous batches from the customer base $\mathcal{I}$ in each batch b.

For customers in the first batch $b=1$, we assign them randomly to the two treatment conditions with a probability of 0.5. For customers subsequent batches, we assign them to the two treatment arms based on the following rule:

$$P_b(W_i=0|\mathbf{X}_i=x)=\frac{\sigma_b^0(x)}{\sigma_b^0(x)+\sigma_b^1(x)}$$
$$P_b(W_i=1|\mathbf{X}_i=x)=\frac{\sigma_b^1(x)}{\sigma_b^0(x)+\sigma_b^1(x)}$$

where $\sigma_b^w(x)$ denotes the standard deviation of the potential outcomes $Y_i(W_i=w)$ estimated from the previous $b-1$ batches. In particular, we estimate customer's response function for the two treatment conditions $(\mathbb{E}[Y_i(0)|\mathbf{X}_i], \mathbb{E}[Y_i(1)|\mathbf{X}_i])$ using two Random Forest models. These Random Forest models are implemented using the `sklearn` package in Python, each consisting of 100 trees with maximum depths not exceeding 10. We estimate $(\sigma_b^0(x), \sigma_b^1(x))$ by computing the standard deviation across the predictions generated by each tree.

We slightly modify the decision estimation phase from the original paper to ensure comparability with our approach. Specifically, we derive the targeting decisions $\hat{\pi}(\mathbf{X}_i)$ using the CATE predictions $\hat{\tau}(\mathbf{X}_i)$ generated by the CATE model, rather than directly estimating them with a policy learning model (e.g. Athey and Wager 2021). This adjustment allows us to eliminate potential differences in targeting performance that may arise from different estimation strategies.[1]

For CATE estimation, we construct a Causal Forest model implemented using the `econML` package in Python, consisting of 100 trees with maximum depths not exceeding 10. Note that the Causal Forest model in `econML` is designed to solve the local moment equation:

$$\mathbb{E}[Y_i-\tau(x)\cdot W_i-B(x)|\mathbf{X}_i=x]=0$$

where $B(x)=\mathbb{E}[Y_i|\mathbf{X}_i=x]$. Therefore, we account for the adaptive nature of the experimental data by substracting the propensity score $P_b(W_i=1|\mathbf{X}_i=x)=\frac{\sigma_b^1(x)}{\sigma_b^0(x)+\sigma_b^1(x)}$ from the actual treatment assignment W_i based on Robinson's Decomposition (Robinson 1988):

$$Y_i-B(\mathbf{X}_i)=\tau(\mathbf{X}_i)(W_i-e(\mathbf{X}_i))+\varepsilon_i$$

[1] Note that the treatment assignment ratio proposed in Kato et al. (2024) remains unaffected by the estimation strategy.

C.1.3. Proposed Approach with Multiple Stages For each experimental size $N_e \in \{1k, 5k, 10k, 20k, 30k\}$, we follow a similar procedure as Waisman et al. (2024) by using equal-sized batches throughout the experiment. In particular, We partition the full sample into $\frac{N_e}{200}$ batches, each containing 200 customers. Customers within each batch b are randomly assigned to the two treatment conditions with a probability of 0.5.

For expected profit loss estimation, as well as for the final CATE estimation, we utilize a Causal Forest model implemented using the `econML` package in Python. This model consists of 100 trees with maximum depths not exceeding 10.

C.1.4. Proposed Approach with Two Stages For each experimental size $N_e \in \{1k, 5k, 10k, 20k, 30k\}$ and each proportion of customers to sample in the first stage $r \in \{0.5, 0.7, 0.9)$, we follow a two-stage sampling approach:

1. In the first stage, we randomly sample $r \cdot N_e$ customers from the customer base $\mathcal{I}$.
2. In the second stage, we select the remaining $(1-r)\cdot N_e$ customers who have the highest expected profit loss estimated from the first stage.
3. Customers sampled in both stages are randomly assigned to the two treatment conditions with a probability of 0.5.

For expected profit loss and final CATE estimation, we employ a Causal Forest model implemented using the `econML` package in Python. This model comprises 100 trees with maximum depths not exceeding 10.

C.2. Additional Results for Different CATE Distributions

In this appendix, we present the results for different CATE distributions, including a bimodal distribution with two equal segments and a bimodal distribution with two unequal segments.

C.2.1. Bimodal Distribution with Two Equal Segments We generate a customer base $\mathcal{I}$ with a bimodal CATE distribution featuring two equal segments according to the following data generating process:

$$Y_i = \tau(X_i) * W_i + X_{i4} * X_{i5} + e_i, \quad e_i \sim \mathcal{N}(0,1)$$

$$\tau(X_i) = X_{i1} * X_{i2} + X_{i3} * (1 - X_{i2})$$

where

$$X_{i1} \sim \mathcal{N}(2, 1.5)$$

$$X_{i3} \sim \mathcal{N}(-4, 1.5)$$

$$X_{i5} \sim \mathcal{N}(0, 1)$$

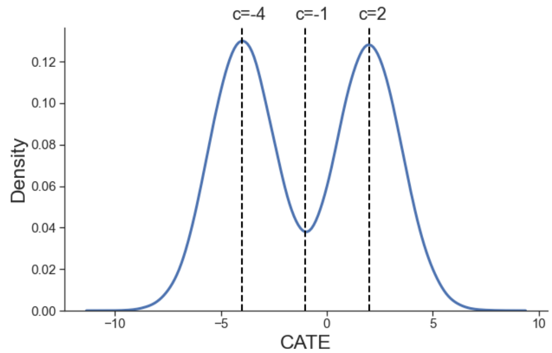
$$X_{ij} \sim Bernoulli(0.5), \quad j \in \{2, 4\}$$

are identically and independently distributed.

In this scenario, we consider three different intervention costs $c \in \{-4, -1, 2\}$ where $c = -4$ and $c = 2$ corresponds to peak of the distribution, and $c = -1$ represents a valley in the CATE distribution. Figure App-1 visualizes the relationship between the CATE distribution and the intervention costs.

Figure App-2 and Figure App-3 shows the proportional profit gaps and profit gaps of the targeting policies learned by different experimental designs across different intervention costs ($c \in \{-4, -1, 2\}$) and experimental sizes ($N_e \in \{1k, 5k, 10k, 20k, 30k\}$) respectively. The results are qualitatively similar to the one with

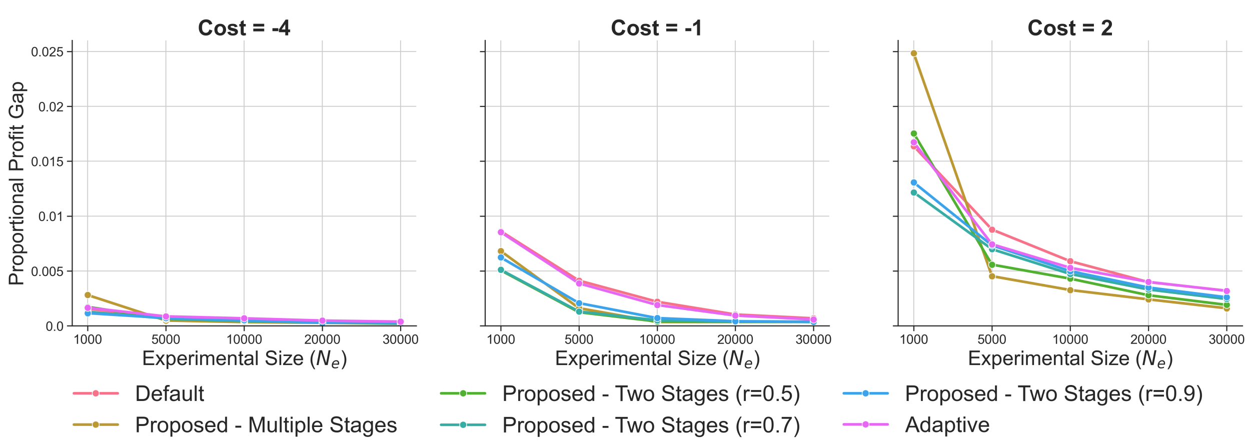
Figure App-1 CATE Distribution and Intervention Costs: Bimodal Distribution with Two Equal Segments



Note. Each dashed line corresponds to a different intervention cost c. $c \in \{-4, 2\}$ aligns with the mode of the CATE distribution. $c = -1$ represents the valley of the CATE distribution.

normally distributed CATEs. In particular, our approach consistently outperforms the default approach and Kato et al. (2024) when the number of customers around the decision threshold is limited (i.e., $c = -1$). Conversely, when the decision threshold aligns with the mode of the distribution (i.e., $c \in \{-4, 2\}$), achieving comparable performance to the two benchmarks requires a larger sample size (in the first stage) to obtain more accurate expected profit loss estimates. [2]

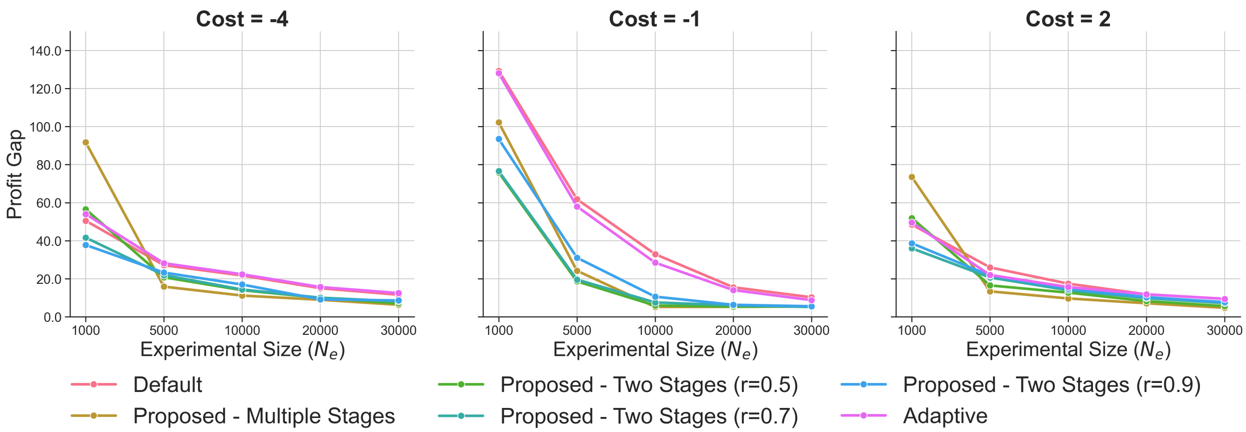
Figure App-2 Proportional Profit Gaps of Different Experimental Designs: Bimodal Distribution with Two Equal Segments



Note. We report the average value of the proportional profit gap across 100 replications. Each line corresponds to an experimental approach.

[2] The difference in the proportional profit gap between $c = -4$ and $c = 2$ is primarily due to the disparity in their denominators. In particular, since $c = -4$ represents a less costly intervention, the incremental profit generated by the optimal policy is greater compared to $c = 2$. As a result, despite having similar numerators, this leads to asymmetry in the proportional profit gaps between the two scenarios.

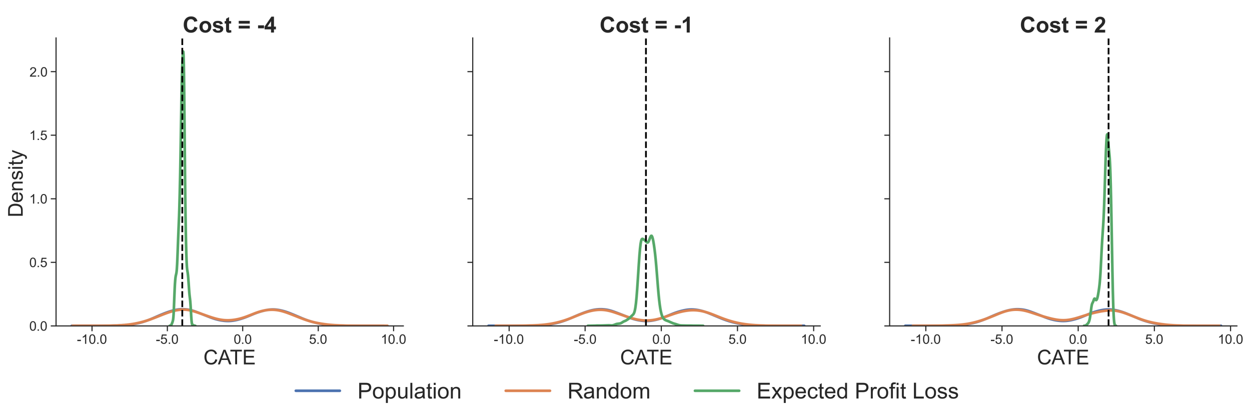
Figure App-3 **Profit Gaps of Different Experimental Designs: Bimodal Distribution with Two Equal Segments**



Note. We report the average value of the profit gap across 100 replications. Each line corresponds to an experimental approach.

Figure App-4 displays the CATE distributions of the customers sampled by various approaches across different intervention costs. The results underscore the effectiveness of our approach in identifying and intensively sampling consequential customers.

Figure App-4 **CATE Distributions of Customers Sampled by Different Approaches: Bimodal Distribution with Two Equal Segments**



Note. Each line corresponds to the CATE distribution of the customers sampled by different approaches. The dashed line represents the intervention cost, which is also the decision threshold. The difference between $c = -4$ and $c = 2$ is mainly driven by the difference in the denominator. In particular, since $c = 2$ exhibits greater cost, the profit of the optimal policy of $c = 2$ is smaller than $c = -4$, leading to a smaller denominator in the proportional profit gap formula. Therefore, the evaluation metric

C.2.2. Bimodal Distribution with Two Unequal Segments We generate a customer base $\mathcal{I}$ with a bimodal CATE distribution featuring two unequal segments according to the following data generating process:

$$Y_i = \tau(X_i) * W_i + X_{i4} * X_{i5} + e_i, \quad e_i \sim \mathcal{N}(0,1)$$

$$\tau(X_i) = X_{i1} * X_{i2} + X_{i3} * (1 - X_{i2})$$

where

$$X_{i1} \sim \mathcal{N}(3,1)$$

$$X_{i3} \sim \mathcal{N}(-2,1)$$

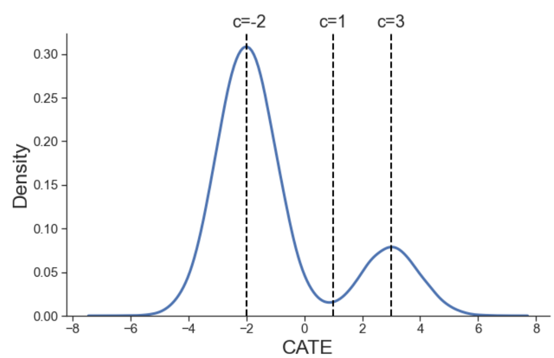
$$X_{i5} \sim \mathcal{N}(0,1)$$

$$X_{ij} \sim Bernoulli(0.5), \quad j \in \{2,4\}$$

are identically and independently distributed.

In this scenario, we examine three different intervention costs $c \in \{-2, 1, 3\}$ where $c = -2$ is at the lager peak, $c = 3$ is at the smaller peak, and $c = 1$ is at the valley of the CATE distribution. Figure App-5 illustrates the relationship between the CATE distribution and the intervention costs.

Figure App-5 CATE Distribution and Intervention Costs: Bimodal Distribution with Two Unequal Segments



Note. Each dashed line corresponds to a different intervention cost c. $c = -2$ aligns with the larger peak and $c = 3$ aligns with the smaller peak of the CATE distribution. $c = 1$ represents the valley of the CATE distribution.

Figure App-6 shows the proportional profit gaps of the targeting policies learned by different experimental designs across different intervention costs ($c \in \{-2, 1, 3\}$) and experimental sizes ($N_e \in \{5k, 10k, 20k, 30k\}$).[3]

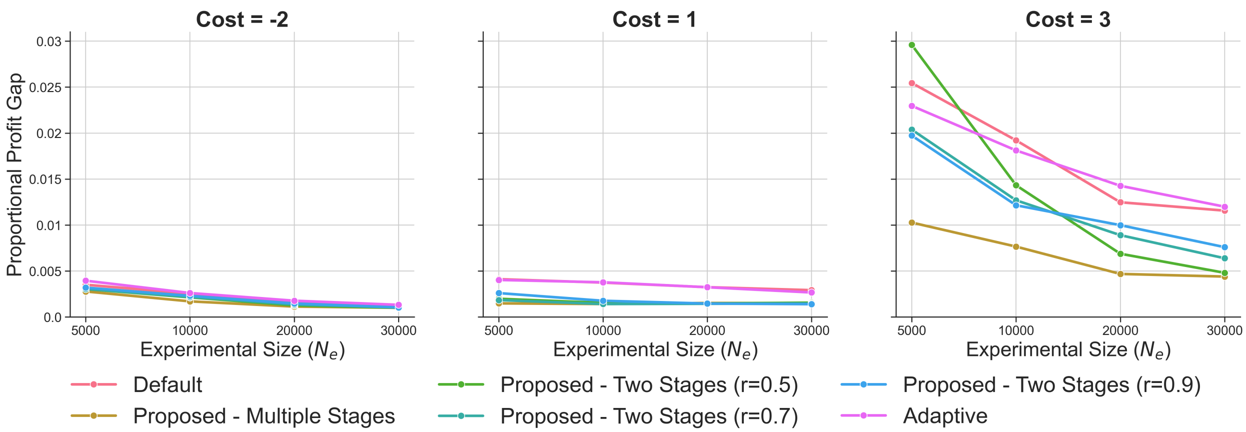
As shown in the graph, our approach generally surpasses the two benchmarks when the cost is at the valley and the smaller peak of the CATE distribution.[4] However, when the sample size is small (i.e., $N_e = 5000$), it is crucial for the firm to carefully select the appropriate deign. Specifically, when the decision threshold is at the smaller peak, the firm should avoid a two-stage design with a limited first-stage sample, as the small initial sample size can lead to substantial expected profit loss estimation errors, hindering the algorithm's

[3] When $c = 1$, due to the scarcity of consequential customers, our approach requires a larger sample size to identify sufficient consequential customers. Therefore, we omit the analysis with $N_e = 1k$ for this case.

[4] When the decision threshold aligns with the valley of the distribution, our approach still outperforms the benchmarks as expected, albeit with smaller magnitude. This reduced improvement occurs because most customers are inconsequential in this case — their prediction errors have minimal impact on targeting profitability — leaving limited room for improvement due to the scarcity of consequential customers.

ability to identify consequential customers. Ultimately, we recommend adopting a two-stage design with a larger proportion of customers sampled in the first stage, especially when the overall sample size is small.

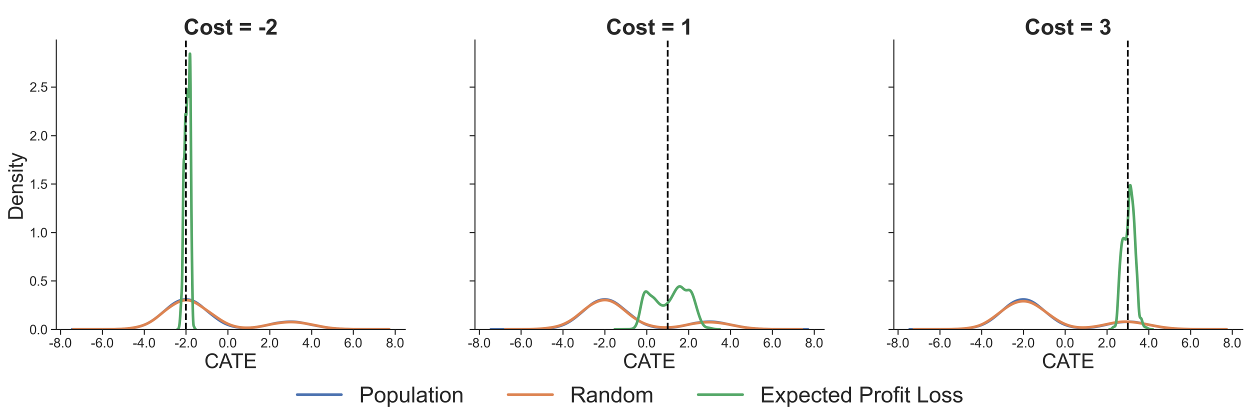
Figure App-6 **Proportional Profit Gaps of Different Experimental Designs: Bimodal Distribution with Two Unequal Segments (Larger Sample Size)**



Note. We report the average value of the proportional profit gap across 100 replications for cases with larger sample size ($N_e > 1000$). Each line corresponds to an experimental approach.

Figure App-7 illustrates the CATE distributions of sampled customers across varying intervention costs for different approaches. The findings highlight how our method effectively identifies and targets consequential customers with heightened intensity.

Figure App-7 **CATE Distributions of Customers Sampled by Different Approaches: Bimodal Distribution with Two Unwqual Segments**



Note. Each line corresponds to the CATE distribution of the customers sampled by different approaches. The dashed line represents the intervention cost, which is also the decision threshold.

Web Appendix D: Further Details of Empirical Application

D.1. Summary Statistics and Randomization Check for Telecommunication Dormant Reactivation Campaign

Table App-2 presents the summary statistics of the pre-treatment covariates for the telecommunication campaign data. Additionally, we perform a weekly randomization check to verify the correct implementation of the randomization process.[5] The results indicate proper randomization, with no significant differences observed between the treatment and control groups across most variables and weeks.

Table App-2 Summary Statistics of Telecommunication Dormant Reactivation Campaign

Variable	Type	Mean	Std.	Median
targeting ratio	Continuous	0.7760	0.1239	0.7985
cancellation usage (7 days)	Continuous	-0.9528	11.0430	0.0
cancellation usage (14 days)	Continuous	-1.9712	22.1519	0.0
cancellation usage (30 days)	Continuous	-4.5623	47.3623	0.0
dawli usage (7 days)	Continuous	0.9518	8.0485	0.0
dawli usage (14 days)	Continuous	2.2245	13.2405	0.0
dawli usage (30 days)	Continuous	10.0042	31.8626	0.0
recharge (7 days)	Continuous	0.4495	5.8295	0.0
recharge (14 days)	Continuous	2.0169	13.9498	0.0
recharge (30 days)	Continuous	16.7776	51.0371	0.0
rental usage (7 days)	Continuous	1.1685	11.7245	0.0
rental usage (14 days)	Continuous	2.4331	23.4193	0.0
rental usage (30 days)	Continuous	6.4194	50.7423	0.0
total usage (7 days)	Continuous	1.4795	9.4835	0.0
total usage (10 days)	Continuous	2.2055	12.2660	0.0
total usage (14 days)	Continuous	3.5324	16.1320	0.0
total usage (30 days)	Continuous	18.2357	45.2000	1.0495
transfer fee usage (7 days)	Continuous	0.0016	0.0519	0.0
transfer fee usage (14 days)	Continuous	0.0038	0.1039	0.0
transfer fee usage (30 days)	Continuous	0.0193	0.3962	0.0
rental usage (7 days)	Continuous	0.3105	3.4168	0.0
rental usage (14 days)	Continuous	0.8421	6.0809	0.0
rental usage (30 days)	Continuous	6.3551	25.4055	0.0

[5] Since customers were randomized on a weekly basis, the randomization check is conducted at the weekly level.

D.2. Summary Statistics and Randomization Check for Starbucks Promotional Campaign

Table App-3 presents the summary statistics of the pre-treatment covariates for the Starbucks data. We also conduct randomization check to verify the correct implementation of the randomization process. The results suggests proper randomization, as there are no significant differences between the treatment and control groups in most of the variables.

Table App-3 Summary Statistics of Starbucks Promotional Campaign

Variable	Type	Mean	Std.	Median
V1 = 0	Discrete	0.1256	–	–
V1 = 1	Discrete	0.3757	–	–
V1 = 2	Discrete	0.3735 –	–	
V1 = 3	Discrete	0.1252	–	–
V2	Continuous	29.9779	5.0009	29.9796
V3	Continuous	0.0	1.0	-0.0395
V4 = 1	Discrete	0.3200 –	–	
V4 = 2	Discrete	0.6800 –	–	
V5 = 1	Discrete	0.1837	–	–
V5 = 2	Discrete	0.3693	–	–
V5 = 3	Discrete	0.3855	–	–
V5 = 4	Discrete	0.0615	–	–
V6 = 1	Discrete	0.2491	–	–
V6 = 2	Discrete	0.2490	–	–
V6 = 3	Discrete	0.2508	–	–
V6 = 4	Discrete	0.2510	–	–
V7 = 1	Discrete	0.2975	–	–
V7 = 2	Discrete	0.7025	–	–

D.3. Implementation Details

D.3.1. Default (A/B Test with Random Sampling) For each experimental size $N_e \in \{10k, 20k, 30k, 40k, 50k\}$, we randomly sample N_e customers from the customer base $\mathcal{I}$. Since the treatment assignments in the original data are properly randomized, we use the treatment assignment from the original data as the final treatment assignment for each sampled customer in our experimentation. For CATE estimation, we construct a Causal Forest model (Wager and Athey 2018) implemented using the `econML` package in Python. This model consists of 300 trees with a maximum depth of 5 to prevent overfitting.

D.3.2. Proposed Approach with Multiple Stages For each experimental size $N_e \in \{10k, 20k, 30k, 40k, 50k\}$, we partition the full sample into $\frac{N_e}{500}$ batches, each containing 500 customers, to streamline the evaluation process. For customers within each batch b, we use the original treatment assignment in the data as their final treatment assignment.

For expected profit loss estimation, as well as for the final CATE estimation, we utilize a Causal Forest model implemented using the `econML` package in Python. This model consists of 300 trees with maximum depths not exceeding 5.

D.3.3. Proposed Approach with Two Stages For each experimental size $N_e \in \{10k, 20k, 30k, 40k, 50k\}$ and each proportion of customers to sample in the first stage $r \in \{0.5, 0.7, 0.9\}$, we follow a two-stage sampling approach:

1. In the first stage, we randomly sample $r \cdot N_e$ customers from the customer base $\mathcal{I}$.
2. In the second stage, we sample the remaining $(1 - r) \cdot N_e$ customers who have the highest expected profit loss estimated from the first stage.
3. For customers sampled in both stages, we use the original treatment assignment in the data as their final treatment assignment.

For expected profit loss and final CATE estimation, we employ a Causal Forest model implemented using the `econML` package in Python. This model comprises 300 trees with maximum depths not exceeding 5.