Multi-Objective Personalization of Marketing Interventions

Omid Rafieian* Cornell University
Anuj Kapoor* IIM Ahmedabad
Amitt Sharma VDO.AI

Abstract

Marketing interventions usually affect multiple outcomes of interest. However, finding an intervention that improves all desired outcomes is often rare, creating a trade-off for managers and decision-makers. In this paper, we develop a multi-objective personalization framework that develops personalized policies to balance multiple objectives at the individual level. We apply our framework to a canonical example of multi-objective conflict between sponsored and organic content consumption outcomes. Partnering with vdo.ai, we conduct a field experiment and randomly assign users to the Skippable/Long and Non-Skippable/Short versions of the same ad. We document substantial substitution between sponsored and organic content consumption: the version that increases sponsored consumption reduces organic consumption. We find that multi-objective personalized policies can significantly improve both sponsored and organic consumption outcomes over single-objective policies. We show that compared to a single-objective policy optimized for organic consumption, there exists a multi-objective policy that increases sponsored consumption by 61% at the expense of only a 4% decrease in organic consumption. Similarly, compared to the single-objective policy optimized for sponsored consumption, there is a multi-objective policy that increases organic consumption by 53% while decreasing sponsored consumption by just 15%.

Keywords: multi-objective personalization, causal inference, machine learning, field experiments, video advertising

*We would like to thank Vaibhav Mishra, Hemank Bajaj, and Kushagra Gupta for excellent research assitance on this project. We thank Sachin Gupta, Jura Liaukonyte, Hema Yoganarasimhan, and Dennis Zhang for detailed comments that have improved the paper. We also thank the participants of the Brown Bag at Cornell University, and marketing seminars at Washington University in St. Louis, Duke, and Wharton for their feedback. Please address all correspondence to: or83@cornell.edu and anujk@iima.ac.in.

1 Introduction

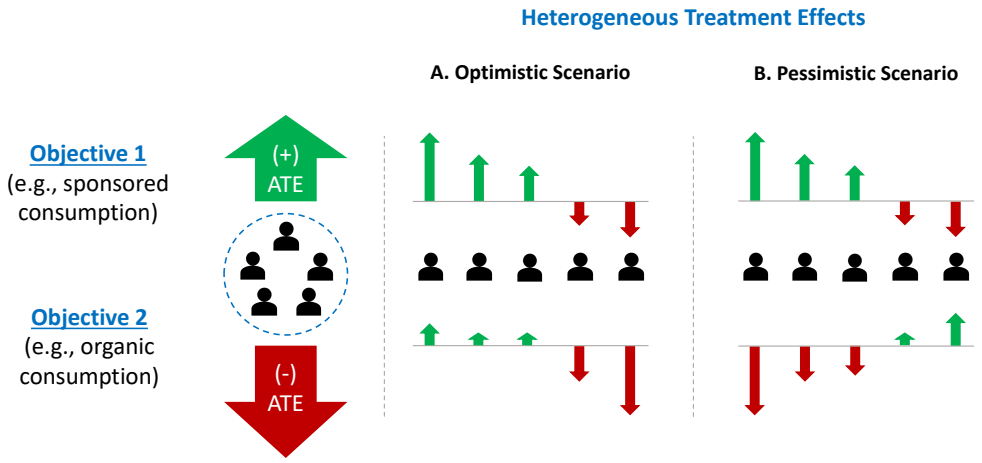
Marketers increasingly rely on experimentation and A/B testing for decision-making. Experimentation identifies the causal effect of each experimental condition relative to a control group, thereby allowing marketers to choose the one with the maximum desired impact. However, most modern marketing problems involve optimizing multiple outcomes. For example, a digital publisher wants to have a higher organic content consumption to increase user engagement as well as a higher sponsored content consumption to achieve higher ad revenues. The main challenge is that these outcomes are often in conflict with each other. That is, actions that improve one desired outcome come at the expense of the other outcome(s), making it unclear what actions to implement. Therefore, the solution to this multi-objective conflict is essential to managers who want to find the right balance between objectives.

Personalization offers a potential solution to this problem. If we break down the heterogeneity in treatment effects at the individual level, we may be able to reduce the multi-objective conflict. Figure 1 illustrates this point using a simple example, where the average treatment effects are in conflict: on average, the treatment causes higher values in one outcome (e.g., organic consumption) but lower values in the other outcome (e.g., sponsored consumption). In an optimistic scenario, the conflict disappears once we look at the treatment effect heterogeneity at the individual level. The treatment effects have the same sign for each individual, so the decision-making is trivial. However, in a pessimistic scenario, the conflict persists even at the individual level: the treatment assignment remains unclear for each individual, as shown in the right panel of Figure 1.

In this paper, we study the problem of multi-objective personalization and examine the value it can create in settings with conflicting objectives. We view the problem through the lens of a platform and address the following sets of questions. The first set of questions is mainly theoretical. How can we design multi-objective personalization algorithms? What are the theoretical guarantees of these algorithms? How can we quantify the performance and the value created by multi-objective personalization algorithms? Next, we turn to a set of empirical questions. What are the gains from using a multi-objective personalization algorithm in a setting with a high degree of substitution? What are the best algorithms in terms of performance? How can managers use data to choose the right policy?

To answer these questions, we face several challenges. First, we need to characterize a specific target

Figure 1. Overview or Multi-Objective Personalization



Note: Arrows show the treatment effects, and colors show the signs: green for positive, red for negative. Each individual is shown with an icon. In the aggregate case, they are grouped together.

or goal for our personalization problem. Unlike the single-objective personalization problem where the expected outcome values are totally ordered, the main challenge in multi-objective personalization is that policies are not directly comparable. This is because one policy can perform better than another policy in one objective but worse in another objective. To address this challenge, we use the concept of Pareto optimality and set our goal as identifying the Pareto frontier of the policy space in terms of multiple objectives, which is the set of policies that are not dominated by any other policy in all objectives. As such, the broad goal of multi-objective personalization is to eliminate all policies that are dominated by at least one feasible policy.

Our second challenge is one of identification: whether we can design multi-objective personalization algorithms that reliably identify the Pareto frontier of the policy space. We combine insights from the recent literature on causal machine learning with that of multi-objective optimization. In particular, we show that the feasible outcome space is convex, which implies that we can identify the complete Pareto frontier using the scalarization approach that maps multiple objectives into a single objective through linear weights. This allows us to design two multi-objective personalization algorithms: (1) *Scalarization with Causal Estimates (`MO-SCE`)*, which uses techniques from heterogeneous treatment effect estimation and proposes policies based on these causal estimates, and (2) *Scalarization with Policy Learning (`MO-SPL`)*, which directly learns the policy that optimizes the weighted objective akin to the literature on policy learning (Swaminathan and Joachims 2015, Athey and Wager 2021). The output of these algorithms is a set of

policies that are on the Pareto frontier of the policy space.

Third, to provide an empirical proof-of-concept for our multi-objective personalization algorithms, we need a setting with a high degree of substitution between outcomes, even at the individual level. Further, for both policy identification and evaluation, we need experimental variation in the treatment assignment. To satisfy these requirements, we partner with the video advertising platform `vdo.ai` and run a field experiment to estimate the effects of ad format on both sponsored and organic content consumption, a canonical example in the marketing literature for two outcomes with a high degree of substitution (Wilbur 2008). In particular, we assign users to three experimental conditions: (1) a Skippable/Long ad for a product, (2) a Non-Skippable/Short ad for the same product, and (3) a No-Ad condition where the user watches the video content without having to watch an ad. The two ad versions are 60-second and 15-second cuts of the same raw footage used by an actual advertiser. We only show pre-roll ads that run before the organic video content and run the experiment for four days on over 50,000 users. The experimental variation in assignment to each ad format allows us to apply our multi-objective personalization algorithms to identify Pareto optimal policies and perform counterfactual policy evaluation for these policies in a confounding-robust manner.

We first estimate the average treatment effect of the Skippable/Long ad version in our study on sponsored and organic consumption metrics relative to the Non-Skippable/Short ad. We measure sponsored consumption as the amount of ad content watched by the user, and organic consumption as the amount of video watched by the user. We find that exposure to the Skippable/Long ad results in 13.5 seconds higher sponsored consumption on average compared to the Non-Skippable/Short ad. Next, we focus on organic consumption as the outcome and show that assignment to the Skippable/Long ad version results in 9.5 percentage points lower organic consumption than the Non-Skippable/Short ad. We then test for the substitution between sponsored and organic consumption. We cannot regress organic consumption on sponsored consumption simply because sponsored consumption is endogenous. However, since our treatment exogenously shifts sponsored consumption, we can use an instrumental variable research design. When we instrument sponsored consumption with the treatment assignment, we find a clear substitution pattern, where every 15 seconds increase in sponsored consumption results in approximately 13 seconds lower organic video consumption, on average.

Next, we explore heterogeneity in treatment effects across observed covariates. We use Causal Forests to estimate the heterogeneity in treatment effects (Wager and Athey 2018). We show substantial variation in the distribution of the Conditional Average Treatment Effect (CATE) on both sponsored and organic consumption. However, the distribution of CATE estimates for each outcome is largely unidirectional: all CATE estimates for sponsored consumption are positive, whereas nearly 97% of CATE estimates for organic consumption are negative. This implies that only for 3.2% of all units, one treatment achieves higher sponsored and organic consumption. Thus, even at the individual level, the platform faces a challenge in finding the right policy that increases both sponsored and organic consumption.

Our multi-objective personalization framework aims to address the substitution between sponsored and organic consumption. Intuitively, multi-objective algorithms identify units whose positive contribution to sponsored consumption outweighs their negative impact on organic consumption and assign these units to the Skippable/Long ad version. We use our multi-objective personalization algorithms and identify a set of policies in each case. To evaluate the performance of these policies, we need a model-free approach, as model-based approaches can favor some algorithms more than others. We turn to the Inverse Propensity Scoring (IPS) estimator proposed by Horvitz and Thompson (1952) that provides an unbiased estimate of the expected outcome under each policy and does not require model-based estimates of the outcomes. To examine how well these algorithms perform, we compare their performance with a group of single-objective personalized policies for each outcome and any random mixing of the two single-objective personalized policies. The gap between our identified Pareto frontier using the multi-objective personalization algorithms and this set of benchmarks quantifies the value multi-objective personalization can create relative to a single-objective personalization algorithm.

Our results reveal a large gap between the Pareto front generated by our algorithms and the set of personalized policies that only use a single objective, indicating that multi-objective personalization creates substantial value in this context. We document that compared to the single-objective personalized policy that only optimizes sponsored consumption, there is a multi-objective personalized policy that increases organic consumption by 53% while only decreasing the sponsored consumption outcome by 15%. We further find larger gains when the platform wants to keep organic consumption high. We show that compared to a single-objective personalized policy that only optimizes organic consumption, there is a policy on the

identified Pareto frontier that improves sponsored consumption by 61% while only reducing organic consumption by 4%. Together, these findings show that multi-objective personalization can create value by substantially improving the performance in one dimension (e.g., organic consumption) without sacrificing the performance in the other dimension (e.g., sponsored consumption).

In sum, our paper offers several contributions to the literature. From a methodological standpoint, we bring insights from multi-objective optimization to the literature on the intersection of machine learning and causal inference and propose a framework for multi-objective personalization. We prove that scalarization identifies the complete Pareto frontier in the outcome space. In particular, we propose two scalarization-based algorithms based on heterogeneous treatment effect estimation and policy learning that can identify the Pareto frontier of policies in terms of multiple objectives. From a managerial perspective, our multi-objective personalization framework can be applied to a wide variety of settings where there is a conflict in treatment effects on multiple outcomes that managers care about. Our framework provides flexibility for managers and decision-makers who want to achieve a certain balance between outcomes by allowing them to evaluate the Pareto frontier a posteriori and select the policy. Substantively, we identify a new area where personalization can create value. The prior literature on personalization has often focused on the richness of covariates to demonstrate the value of personalization where we can differentiate between users. This work uses the variation in multiple outcomes to develop better personalized policies. In particular, we show substantial gains in one outcome without sacrificing another in a domain with a high degree of substitution. More specifically, our work contributes to the literature on the interplay between sponsored and organic consumption. We highlight an important trade-off in the effect of ad format on these outcomes and show that advertising platforms can incorporate a multi-objective approach when designing their ad allocation policies.

2 Related Literature

First, our paper relates to the literature on personalization. User tracking and algorithmic decision-making allow digital platforms to easily implement personalized policies at scale (Lambrecht and Tucker 2019). Recent methodological developments in this literature have brought a causal lens to machine learning algorithms that have been traditionally used for personalization tasks (Swaminathan and Joachims 2015, Shalit et al. 2017, Wager and Athey 2018, Nie and Wager 2021, Athey and Wager 2021). Applied papers in this

domain have documented the gains from personalization in a variety of domains, such as marketing mix personalization for optimal admission and scholarship outcomes (Belloni et al. 2012), incentives in churn management problems (Ascarza 2018), promotional offers in retail settings (Simester et al. 2020a,b), allocation and sequencing of mobile in-app advertising (Rafieian and Yoganarasimhan 2021, Rafieian 2023), length of free trial in software as service industry (Yoganarasimhan et al. 2022), and product versioning in music streaming platforms (Goli et al. 2022). The key insight in this series of work is that having a fine-grained set of pre-treatment variables helps differentiate between users, thereby creating value by assigning users to the right policy. We extend this literature by proposing a multi-objective personalization framework that allows firms to identify Pareto optimal policies that generate considerable gains in many dimensions. In particular, we show that even in a context where single-objective personalized policies offer limited differentiation, a multi-objective personalization approach can create substantial value by differentiating between users based on the magnitudes of treatment effects and substitutability between outcomes at the individual level. Our generic and flexible framework makes it applicable to many marketing and non-marketing problems where the manager needs to optimize more than one objective.

Second, our paper relates to the literature on multi-objective optimization (Marler and Arora 2004). This literature has proposed a series of algorithms to deal with multi-objective optimization problems, ranging from scalarization techniques (Miettinen and Mäkelä 2002) to genetic algorithms (Deb et al. 2002). More closely related to our paper is the stream of literature that considers discrete policies that map the covariate vector to a specific treatment condition, such as the literature on multi-objective contextual bandits (Tekin and Turğay 2018, Turgay et al. 2018, Wang et al. 2022, 2023) and multi-objective reinforcement learning (Roijers et al. 2013, Van Moffaert and Nowé 2014, Abdolmaleki et al. 2020). Our paper builds on the foundations laid out in multi-objective optimization literature and contributes to this stream of work in two ways. First, we bring a causal lens to the problem and directly incorporate the Conditional Average Treatment Effect (CATE) estimates to the problem. Second, we theoretically show that the scalarization method can fully recover the Pareto frontier in this class of problems as the expected outcome space under feasible policies is convex.

Our paper also relates to the literature on the interplay between sponsored and organic content (Sun and Zhu 2013). With advancements in ad measurements, the literature on TV advertising has documented

a phenomenon called "zapping", which is the practice of switching channels during commercial breaks (Zufryden et al. 1993, Danaher 1995, Siddarth and Chattopadhyay 1998). Since then, a series of papers have examined different aspects of ad avoidance by deriving the equilibrium properties in the market users are averse to ads (Anderson and Coate 2005, Dukes et al. 2022), quantifying the audience loss caused by ad avoidance (Wilbur 2008, Schweidel and Moe 2016, Rajaram et al. 2023), predicting ad avoidance based on the past consumption of the product (Tuchman et al. 2018), linking ad avoidance to sales (Bronnenberg et al. 2010, Deng and Mela 2018), and proposing market design solutions to account for audience externalities (Wilbur et al. 2013).[1] We contribute to this stream of work by causally identifying a substitution pattern between sponsored and organic content consumption. Importantly, we show how platforms can use personalization to efficiently exploit this substitution pattern and achieve desired outcomes in both Ad Consumption (sponsored) and Video Consumption (organic).[2]

3 Multi-Objective Personalization

In this section, we present our framework for multi-objective personalization. We first formally present some preliminaries from the literature on causal inference and multi-objective optimization and clearly define our problem in §3.1. We then present our theoretical results on the multi-objective personalization problem and develop two algorithms that help identify multi-objective policies in §3.2. Finally, in §3.3, we present an intuitive measure of performance for the set of policies generated that quantifies the value from multi-objective personalization and helps with model selection.

3.1 Problem Definition

At a high level, multi-objective personalization entails developing a personalized policy that performs well in terms of multiple objectives. To characterize this problem, we need to first define what we mean by a *personalized policy* and *performance*. Let X and W denote the covariates and treatment status, respectively. To further formalize the problem, we let $\mathcal{X}$ and $\mathcal{W}$ denote the support for covariates and treatment. For example, we have $\mathcal{W} = \{0, 1\}$ in a binary treatment context. For notational simplicity, we follow the common norm in the literature and study the case for the binary treatment, which is also consistent with our

[1] Please see Wilbur (2016) for a great summary of the ad avoidance literature.

[2] We use sponsored consumption interchangeably with Ad Consumption, and organic consumption interchangeably with Video Consumption throughout the paper.

empirical application. However, it is easy to extend our algorithms to multiple treatment settings. In a multi-objective personalization problem, we have K different outcomes denoted by $\{(Y^{(1)}, Y^{(2)}, \cdots, Y^{(K)})\}$. For each outcome $Y^{(j)}$, we define the set of potential outcomes as $\{Y^{(j)}(w)\}_{w \in \mathcal{W}}$. With these model preliminaries defined, we can characterize the key concepts in our problem. We start with the definition of a policy π as follows:

Definition 1. *A policy $\pi : \mathcal{X} \to [0, 1]$ is a mapping from the covariate space to probability values that determines the probability of treatment assignment for any observation based on its vector of characteristics X_i.*

Naturally, the definition above implies that the probability of assignment to control is $1 - \pi(X_i)$. For deterministic policies, $\pi(\cdot)$ only takes values in $\{0, 1\}$. Finding a personalized policy is a search over infinitely many options. To perform this search effectively, we need a performance measure tied to our multiple outcomes. For example, in our empirical context, we want to know how each policy performs in terms of sponsored and organic consumption. We define these performance measures as follows:

Definition 2. *For each outcome $Y^{(j)}$, we define the performance of the policy in terms of that outcome as a mapping $\rho_j : \Pi \to \mathbb{R}$, where Π is the space of all possible policies. This indicates that for a policy π, the performance in terms of outcome $Y^{(j)}$ is characterized by $\rho_j(\pi)$. We can formally define this term as follows:*

$$\rho_j(\pi) = \mathbb{E}\left[\pi(X_i)Y_i^{(j)}(1) + (1 - \pi(X_i))\, Y_i^{(j)}(0)\right], \tag{1}$$

where the expectation is taken over the joint distribution of the covariates. Intuitively, $\rho_j(\pi)$ is the expected value of the outcome $Y^{(j)}$ if we implement policy π.

In a multi-objective personalization problem, we need to consider performance metrics for all outcomes of interest. As a result, comparing two policies is more challenging in a multi-objective case compared to a single-objective case, where there is one objective with a clear order. For example, in the context of our problem, we can consider two objectives $\rho_s(\pi)$ and $\rho_o(\pi)$ that are defined for the sponsored consumption and organic consumption outcomes, respectively. If there are two policies π_1 and π_2 such that $\rho_s(\pi_1) > \rho_s(\pi_2)$ and $\rho_o(\pi_1) < \rho_o(\pi_2)$, it is not clear which one the platform must choose. However, if there are two policies π_1 and π_2, such that $\rho_s(\pi_1) > \rho_s(\pi_2)$ and $\rho_o(\pi_1) > \rho_o(\pi_2)$, we can conclude that policy π_2

is dominated by π_1 with respect to both objectives. This type of comparison immediately brings us to the notion of Pareto optimality, where the Pareto frontier of the policy space is the set of policies that are non-dominated by any other policies. To this end, we define the main goal of multi-objective personalization as follows:

Definition 3. *Suppose that there is a manager who wants to optimize multiple outcomes $Y^{(1)}$, $Y^{(2)}$, ..., $Y^{(K)}$. Our goal is to find a set of policies Π_f that are Pareto optimal in terms of objectives $\rho_1(\pi)$, $\rho_2(\pi)$, ..., $\rho_K(\pi)$. That is, for each $\pi \in \Pi_f$, there is no other policy π' in the space of policies such that we have $\rho_j(\pi') > \rho_j(\pi)$, for every j.*

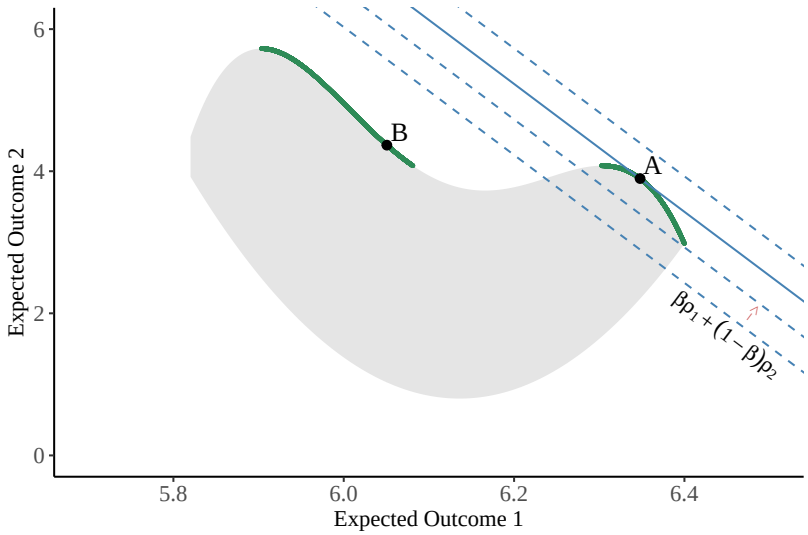
Intuitively, our goal is to eliminate dominated policies and present the manager with a set of Pareto optimal policies Π_f. In the next section, we discuss how we can design algorithms to identify these Pareto optimal policies.

3.2 Algorithms for Multi-Objective Personalization

Fundamentally, the problem in Definition 3 can be viewed as a multi-objective optimization problem. The literature on multi-objective optimization offers many solutions to this problem, given the setting. Much of this literature focuses on the problem with a set of continuous control variables set by the decision-maker that are linked to multiple notions of reward or objective (Marler and Arora 2004). For example, a driver can set continuous variables speed and total passenger weight to optimize the travel time and fuel cost. Multi-objective personalization problem involves finding a complex individual-specific policy that performs well on multiple objectives. As such, our problem is more closely related to the literature on multi-objective contextual bandits (Drugan and Nowe 2013, Tekin and Turğay 2018, Turgay et al. 2018, Wang et al. 2022, 2023) and multi-objective reinforcement learning (Roijers et al. 2013, Van Moffaert and Nowé 2014, Abdolmaleki et al. 2020). Most of these papers focus on regret bounds in an online setting with exploration or use a very specific type of multi-objective optimization problem. Our work considers a general personalization problem in an offline setting and proposes algorithms that can provably identify the complete Pareto frontier with statistical guarantees. A novel aspect of our work is in bringing a causal lens to this problem, which allows us to propose solutions that are generalizable and robust to confounding.

One of the most popular solutions to the multi-objective optimization problem is linear scalarization, whereby we map multiple objectives into a single objective using linear weights that come from the proba-

Figure 2. Identification of Pareto Optimal Points Using Linear Scalarization



bility simplex. That is, for any set of non-negative weights $\beta \in \mathbb{R}_+^K$ such that $\sum_{j=1}^K \beta_j = 1$, we maximize $\sum_{j=1}^K \beta_j \rho_j(\pi)$. It is easy to show for any given set of weights, the optimal solution $\sum_{j=1}^K \beta_j \rho_j(\pi)$ is Pareto optimal.[3] However, there is no guarantee that we can identify the full Pareto frontier by enumerating all possible weights. Figure 2 helps illustrate this point in a setting with two objectives. Geometrically, linear scalarization finds the optimal solution by shifting the lines in the direction illustrated. As such, it can only identify Pareto optimal solutions that are on the convex hull of the outcome space (e.g., point A in Figure 2) but fail to identify Pareto optimal solutions that are not on the convex hull of the outcome space (e.g., point B in Figure 2). Thus, in the general case, linear scalarization methods cannot identify the complete Pareto frontier.

Our multi-objective personalization problem differs from the general class of multi-objective problems as we can formally prove that the outcome space $\{(\rho_1(\pi), \ldots, \rho_K(\pi))\}_{\pi \in \Pi}$ is convex. The intuition for this result is as follows: for any two points in the outcome space, we can find the set of policies that cover the line between these two points using a mixed strategy policy that probabilistically uses these two policies. Thus, we can write the following proposition:

Proposition 1. *Let* $\mathcal{B} = \{\beta \mid \beta \in \mathbb{R}_+^K, \sum_{j=1}^K \beta_j = 1\}$ *denote the full set of linear weights. For any* $\beta \in \mathcal{B}$*, we define* $\pi_\beta^S = \operatorname{argmax} \sum_{j=1}^K \beta_j \rho_j(\pi)$*, which is the solution to single objective problem with weights* β*. The full set of policies identified by linear scalarization, defined as* $\Pi^S = \{\pi_\beta^S\}_{\beta \in \mathcal{B}}$*, is the complete Pareto*

[3] We can prove it by contradiction: if there is a policy π' that dominates π in all objectives, then it is a contradiction that π is the optimal solution to the scalarized objective.

frontier of the outcome space, i.e., $\Pi^S = \Pi_f$.

Proof. See Appendix A.1 for the proof. □

This proposition theoretically shows that we can use linear scalarization to identify the complete Pareto frontier if we know the function $\rho_j(\cdot)$ for any k. The empirical challenge is that we do not know $\rho_j(\cdot)$ and we need to estimate it from the data. For any set of weights $\beta \in \mathcal{B}$, we can write the scalarized optimization problem as follows:

$$\pi_\beta^S = \text{argmax} \sum_{k=1}^{K} \beta_k \rho_k(\pi) \tag{2}$$

In the following sections, we propose two algorithms that take data as the input and use linear scalarization to empirically identify the Pareto frontier. Before we proceed with our empirical algorithms, we make a key assumption:

Assumption 1. *The following four conditions hold for data set* $\mathcal{D} = \{(X_i, W_i, Y_i^{(1)}, \ldots, Y_i^{(K)})\}_{i=1}^N$*: (1) Stable Unit Treatment Value Assumption (SUTVA), which states that potential outcomes for one unit are not influenced by other units' treatment assignment and there is only one version for each treatment, (2) Unconfoundedness assumption, which states that the treatment assignment is independent of potential outcomes given observed covariates, (3) Overlap, which states that the treatment assignment is probabilistic, and (4) No Distribution Shift assumption, which indicates that the joint distribution of covariates and potential outcomes is fixed.*

The four conditions in the assumption above are all standard causal inference assumptions that are largely used in the personalization literature (Rafieian and Yoganarasimhan 2023). As we will discuss later, these assumptions allow us to empirically identify the Pareto frontier using our algorithms. It is worth emphasizing that since we do not know the function $\rho_j(\cdot)$, our task at hand is one of statistical estimation, which naturally involves some uncertainty and error. As such, Pareto frontiers identified under different algorithms can differ due to the randomness in the data. Súkeník and Lampert (2022) show that one could build generalization error bounds and excess bounds for scalarized objectives like in Equation (2) in order to bound the difference between the *empirically* Pareto optimal set of policies obtained by algorithms and the *truly* Pareto optimal set of policies. Therefore, the statistical guarantees our algorithms have for the single-objective case will conveniently transfer to the multi-objective case.

3.2.1 Scalarization with Causal Estimates

In this section, we present our first algorithm to empirically identify the Pareto frontier, which uses causal estimates to simplify the optimization in Equation (2). We first define our key causal estimand as follows:

Definition 4. *For any outcome $Y^{(j)}$, the Conditional Average Treatment Effect (CATE) is denoted by $\tau_k(\cdot)$ and defined as follows:*

$$\tau_k(x) = \mathbb{E}\left[Y_i^{(j)}(1) - Y_i^{(j)}(0) \mid X_i = x\right] \tag{3}$$

Intuitively, the CATE estimate measures the treatment effect conditional on a certain value of the covariates. The prior literature in the intersection of causal inference and machine learning offers a host of methods that estimate CATE under the set of assumptions presented in Assumption 1 (Shalit et al. 2017, Wager and Athey 2018). These estimators have desirable statistical properties such as consistency and unbiasedness. Using the potential outcomes framework and the CATE definition, we can write the following proposition about the optimal personalized policy given any β:

Proposition 2. *For any $\beta \in \mathcal{B}$, the optimal solution to the objective function in Equation (2) is as follows:*

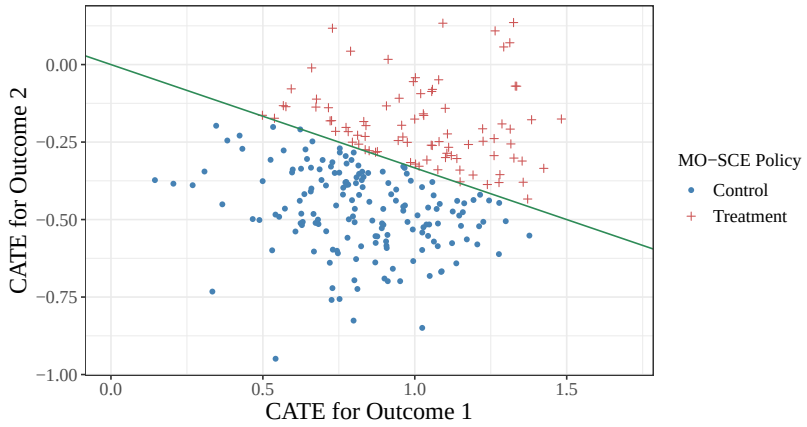
$$\pi_\beta(x) = \mathbb{1}(\sum_{j=1}^{K} \beta_j \tau_j(x) \geq 0), \tag{4}$$

where $\tau_j(\cdot)$ is the CATE on outcome $Y^{(j)}$.

Proof. Please see Web Appendix A.2. □

The intuition behind Proposition 2 is the following: the only policy-variant element of ρ_j is τ_j. Hence, we can simplify the policy optimization in Equation (2) and write it in terms of CATEs. This also simplifies the empirical task at hand, as we only need to estimate CATEs. Figure 3 illustrates how Proposition 2 works by using a simple example with two objectives, where the line for a given β determines which observations should receive the treatment or control. For other weights, the line shown will rotate in a certain direction and identify different policies. Our Scalarization with Causal Estimates algorithm precisely does that: it forms the union of all policies generated by different values of $\beta \in \mathcal{B}$.

Figure 3. An Illustration of the Assignment Policy in Scalarization with Causal Estimates



Note: Each point shows the CATE on Outcomes 1 and 2 for a data point. The figure illustrates the Scalarization with Causal Estimates policy for $\beta = (0.25, 0.75)$, so points above the line $0.25\tau_1 + 0.75\tau_2$ are assigned to treatment (red plus), and points below that line to the control condition (blue circle).

We present our algorithm for Scalarization with Causal Estimates in Algorithm 1. The algorithm takes CATE estimates as inputs and generates different policies corresponding to each element of the set $\mathcal{B}$. As shown in this algorithm, for any set of weights $\beta \in \mathcal{B}$, we use CATE estimates to obtain a personalized policy. The output is a set of policies that is the identified Pareto frontier using this algorithm. In Appendix B.1, we present a very simple illustrative example with three data points.

Algorithm 1 Scalarization with Causal Estimates (`MO-SCE`)

Input: $\{(\hat{\tau}_1(X_i), \hat{\tau}_2(X_i), \ldots, \hat{\tau}_K(X_i))\}_{i=1}^N$, $\mathcal{B}$
Output: $\Pi^{\text{MO-SCE}}$

1: **for** $\forall(\beta_1, \beta_2, \ldots, \beta_K) \in \mathcal{B}$ **do**
2: $\quad \pi_\beta^{SCE}(x) \leftarrow \mathbb{1}\left(\sum_{j=1}^K \beta_j \hat{\tau}_j(x) \geq 0\right)$
3: **end for**
4: $\Pi^{\text{MO-SCE}} \leftarrow \bigcup_{\beta \in \mathcal{B}} \pi_\beta^{\text{MO-SCE}}(x)$

The overall logic of Algorithm 1 is to first estimate CATEs and then feed these estimates to the algorithm as inputs to solve for Equation (4). Because the main uncertain piece in Equation (4) is the set of CATE estimates, the statistical properties of the algorithm depend on the statistical properties of the CATE estimator. Under Assumption 1, we know that a host of methods can produce consistent and unbiased estimates of CATE, which offers statistical guarantees for Algorithm 1. More specifically, the Scalarization with Causal Estimates algorithm performs well when the CATE estimates are more accurate.

3.2.2 Scalarization with Policy Learning

An alternative empirical approach to identify policies that optimize the objective function in Equation (2) is to use direct policy learning (Swaminathan and Joachims 2015, Athey and Wager 2021). In a single-objective setting where we only want to optimize $\rho_j(\cdot)$, this approach uses an unbiased estimate of $\rho_j(\cdot)$ to form an objective function. It then identifies the policy π from a certain policy class Π^* that optimizes the objective function. In this section, we want to define this approach for the multi-objective setting as in Equation (2). To do so, we define an estimator for $\rho_j(\cdot)$ as follows:

Definition 5. *Let $e(X_i)$ denote the propensity score for treatment. For any data $\mathcal{D}$ with N observations, the Inverse Propensity Scoring (IPS) estimator for outcome $Y^{(j)}$ under policy π is defined as follows:*

$$\hat{\rho}_j^{\text{IPS}}(\pi; \mathcal{D}) = \frac{1}{N} \sum_{i \in \mathcal{D}} \left(\frac{\pi(X_i) W_i}{e(X_i)} + \frac{(1 - \pi(X_i))(1 - W_i)}{1 - e(X_i)} \right) Y_i^{(j)}. \tag{5}$$

The IPS estimator was first proposed by Horvitz and Thompson (1952) and then used widely for counterfactual policy evaluation in the personalization literature (Rafieian and Yoganarasimhan 2023). Importantly, the IPS estimator $\hat{\rho}_k^{IPS}(\cdot)$ is an unbiased estimator for $\rho_j(\cdot)$. As such, we can define the empirical equivalent of Equation (2) by plugging in $\hat{\rho}_j^{IPS}(\cdot)$ for every $\rho_j(\cdot)$. In the following proposition, we show what optimizing this empirical target is equivalent to:

Proposition 3. *Suppose we have data $\mathcal{D} = \{(X_i, W_i, e(X_i), Y_i^{(1)}, \ldots, Y_i^{(K)})\}_i$. The variable Γ_i is defined at the individual level as follows:*

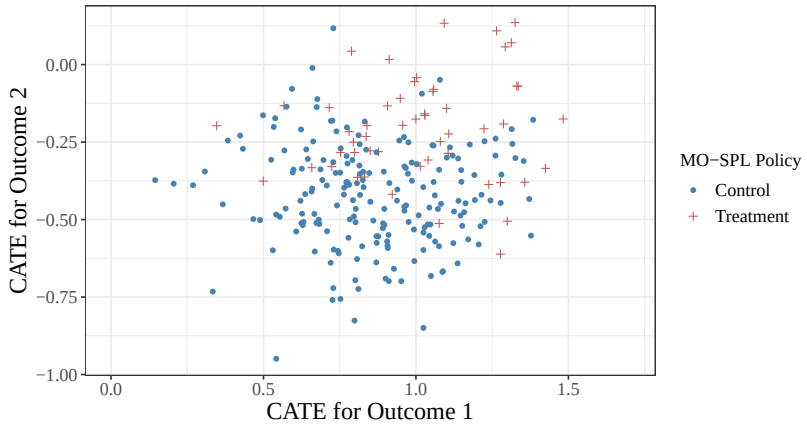
$$\Gamma_i = \left(\frac{W_i}{e(X_i)} - \frac{1 - W_i}{1 - e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \tag{6}$$

For a given deterministic policy class Π^*, the solution to the $\text{argmax}_{\pi \in \Pi} \sum_{j=1}^{K} \beta_j \hat{\rho}_j^{\text{IPS}}(\pi)$ will be the solution to a weighted classification problem with $L_i = \text{sgn}(\Gamma_i)$ as the label and $|\Gamma_i|$ as weights.

Proof. Please see Web Appendix A.3. □

This proposition indicates that for any $\beta \in \mathcal{B}$, we can directly learn policy π_β^S in Equation (2), using any off-the-shelf classification algorithm. The intuition behind this is the following: a more positive Γ_i

Figure 4. An Illustration of the Assignment Policy in Scalarization with Policy Learning



Note: Each point shows the CATE on Outcomes 1 and 2 for a data point. The figure illustrates the Scalarization with Policy Learning policy for $\beta = (0.25, 0.75)$. The treatment (red plus) and control (blue circle) assignments are determined based on the weighted classification task.

indicates greater gains when assigned to the treatment, whereas a more negative one indicates greater gains when assigned to the control condition. A weighted classification tries to learn the best assignments. Figure 4 shows an example of treatment assignment under the policy learning approach. Unlike Figure 3, the policy is not linear in CATE estimates, but follows the same logic: points with higher CATEs on both dimensions are more likely to receive the treatment.[4]

We present the algorithm for multi-objective personalization that uses the policy learning approach in Algorithm 2. Like Algorithm 1, we obtain an estimate of policy π^S_β for any $\beta \in \mathcal{B}$. The resulting output is a set of size $|\mathcal{B}|$, which is the empirically identified Pareto frontier using the policy learning algorithm. We present a very simple illustrative example with three data points in Appendix B.2.

Like our Scalarization with Causal Estimates algorithm, we require the standard assumptions of causal inference presented in Assumption 1 for identification. However, one key advantage of this approach is that even without theoretical guarantees on the $1/\sqrt{N}$-rate estimation of CATEs, we can identify the optimal policy at a $1/\sqrt{N}$-rate regret over the policy class Π if we know the propensity scores and the VC dimension of Π is finite (Athey and Wager 2021). Another advantage of this algorithm is that we can directly give

[4]It is worth emphasizing that the reason why the two algorithms identify different policies is the randomness involved in the empirical application. Further, unlike the Scalarization with Causal Estimates policy that only uses CATE estimates, the Scalarization with Policy Learning algorithm uses other covariate inputs and map them to policies, which induces some natural differences in the policy assignment. Finally, some individual-level differences between policies are less consequential for the final estimate of the expected outcome under a policy across all users, especially when assignment to either policy does not change the objective. For example, for points around the line in Figure 3, the assignment to either policy does not drastically shift the objective.

Algorithm 2 Scalarization with Policy Learning (`MO-SPL`)

Input: $\{X_i, W_i, e(X_i), Y_i^{(1)}, \ldots, Y_i^{(K)}\}_i, \mathcal{B}$
Output: $\Pi^{\text{MO-SPL}}$

for $\forall(\beta_1, \beta_2, \ldots, \beta_K) \in \mathcal{B}$ **do**
  $\Gamma_i \leftarrow (\frac{W_i}{e(X_i)} - \frac{1-W_i}{1-e(X_i)})(\sum_{j=1}^{K} \beta_j Y_i^{(j)})$
  $\pi_\beta^{SPL} \leftarrow \text{argmax}_{\pi \in \Pi}(2\pi(x) - 1)\Gamma_i$ ▷ Classification: $\text{sign}(\Gamma_i)$ as labels and $|\Gamma_i|$ as weights
end for
$\Pi^{\text{MO-SPL}} \leftarrow \bigcup_{\beta \in \mathcal{B}} \pi_\beta^{\text{MO-SPL}}(x)$

CATE estimates as inputs in addition to the data $\mathcal{D}$, facilitating the policy learning process. On the other hand, the main drawback of this approach is for cases where propensity scores are small, which creates a large variance in the Γ weights. In order to overcome this challenge, the natural solution in the literature is to use a doubly robust scoring as in Athey and Wager (2021).

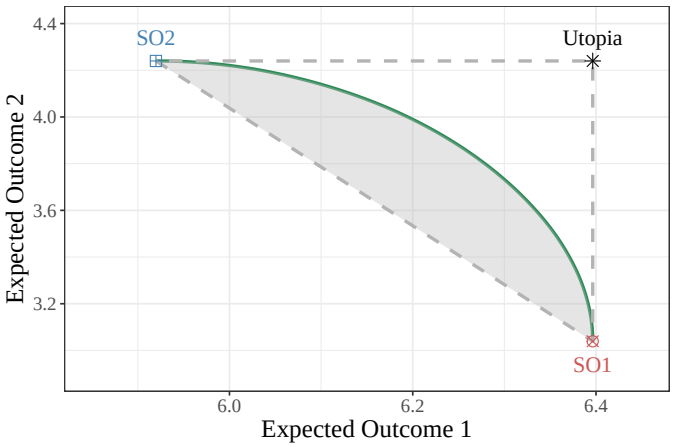
3.3 Evaluation of Set of Policies

Suppose we have two algorithms A and B that identify policies π_A and π_B in a single-objective personalization problem. If we have an evaluator $\hat{\rho}(\cdot)$, we can compare the performance of π_A and π_B and see which one performs better. In a multi-objective personalization setting, however, algorithms A and B identify a set of policies Π_A and Π_B, respectively. As such, the comparison is not trivial as we are dealing with sets of policies. In this section, we use the theoretical structure of our problem and propose a simple measure that allows us to compare the performance of two sets.

We use a 2-dimensional example in Figure 5 to illustrate our proposed measure. Each point in this figure corresponds to the expected outcomes under a policy. We see the expected outcomes under two single-objective personalized policies, SO1 and SO2. Given that the outcome space is convex, we know that the weakest identified Pareto frontier would be the line between these two single objective policies, as one could achieve any point on this line by using a mixed strategy policy that uses either single-objective personalized policy with a certain probability. The line between the two single-objective personalized policies thus serves as a reasonable benchmark for multi-objective personalization. On the other hand, the best-case scenario for multi-objective personalization is to achieve a *Utopia* point shown in Figure 5 that takes the optimal expected outcome in each dimension. The Pareto frontier will naturally be a curve within the triangle between single-objective policies and the utopia point.

Intuitively, the farther the Pareto frontier gets from the single-objective line and the closer it gets to the

Figure 5. A Visual Illustration of Multi-Objective Policy Evaluation



Note: SO1 and SO2 refer to the single-objective personalized policies for outcomes 1 and 2, respectively. A probability mixing between the two covers the line between SO1 and SO2. The Utopia point represent the optimal values in each dimension. The green curve is the Pareto frontier under a multi-objective algorithm. The Covered Area Proportion (CAP) measure is the proportion of shaded area to the triangle between SO1, SO2, and Utopia.

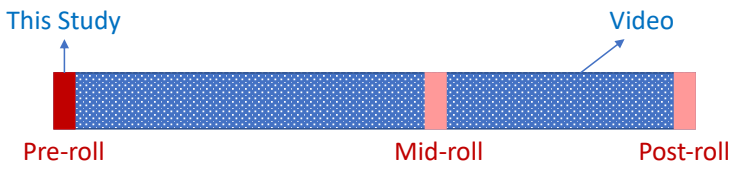
utopia, the greater the value created by multi-objective personalization. Therefore, the area between the Pareto frontier and the line between single-objective personalized policies indicates the performance of the set of Pareto optimal policies. To normalize this measure, we use the proportion of this area to the total area of the triangle as our main measure and call it the *Covered Area Proportion (CAP)*. We can easily extend this notion to multi-dimensional settings and evaluate the set of Pareto optimal policies.

An important benefit of using the CAP measure is in model selection. As discussed earlier, the `MO-SCE` performs well when we can accurately estimate the CATEs from data. On the other hand, the `MO-SPL` algorithm does not require accurate CATE estimates, but requires accurate and not too small propensity weights. In many cases, it is not easy to verify these conditions ex-ante. Therefore, researchers can use the CAP measure to choose the best-performing algorithm for empirical identification of the Pareto frontier.

4 Empirical Application

In our empirical application, our primary goal is to provide a proof-of-concept for our multi-objective personalization algorithms. As such, we need a setting where there is a high degree of conflict between the treatment effects on multiple outcomes of interest. Further, we need to have an experiment so we can reliably measure the performance of policies proposed by our algorithms. To meet these criteria, we collaborate with the video advertising platform `vdo.ai` and conduct a field experiment to measure the

Figure 6. Different Types of Video Ads



impact of ad format on both sponsored and organic content consumption, outcomes that exhibit a high degree of substitution (Wilbur 2008).

Before we proceed with the details of our empirical application, we stress that the main purpose of our empirical application is to provide the right setting to test our multi-objective personalization framework and acknowledge that obtaining generalizable insights from our experiment is challenging, given the heterogeneity in the effects of interest and the specificity of our empirical context.

4.1 Experiment

The application setting of our study is the video advertising industry. We partner with the company `vdo.ai`, which is based in India and the US and provides video services to publishers worldwide. Since its launch, `vdo.ai` has attracted many large and medium-sized media publishers who use the company's technology to serve organic and sponsored video content on their web pages. As a form of monetization, `vdo.ai` places video ads at different parts of an organic video. The sponsored content in our study is in the form of video ads. The video ads are generally of three types: (1) pre-roll ads that are placed prior to the start of the video, (2) mid-roll ads that are placed in the middle of the video, and (3) post-roll ads that are placed after the content video has finished playing. Figure 6 visualizes these different types of ads. In our experiment, we only focus on the pre-roll ads that are shown before the organic content starts.

Unlike video streaming platforms such as YouTube, the organic video content in our context is not generally the primary reason a user is visiting a web page. For example, the organic video content may be placed in a news article a user wants to read. This is a common practice among media publishers such as `cnn.com` or `espn.com` who display organic and sponsored content within the news article. As the main facilitator of sponsored and organic video content, `vdo.ai` wants a higher sponsored consumption as it is directly linked to ad revenues[5], as well as a higher organic consumption as it relates to user engagement

[5] Advertisers care about the ad effect on conversion outcomes, which determines their willingness to pay for units of sponsored consumption. In our analysis, we assume a fixed price per unit of sponsored consumption. It is easy to relax this assumption and directly optimize for ad revenues.

and attracts more content creators, which helps the platform generate more advertising opportunities in the short and long run.

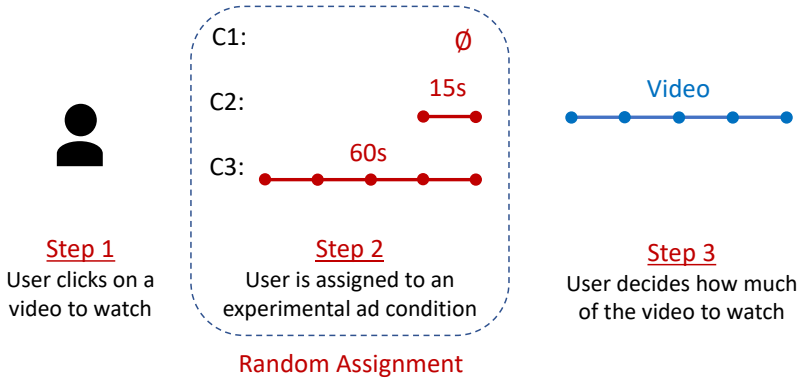
The platform uses two different inventories of ad impressions to allocate video ads. In the first inventory, a second-price auction determines which ad will be placed in an impression. That is, advertisers participate in an auction, and the impression will be awarded to the ad with the highest bid or willingness to pay. The second inventory is an unsold impression inventory used for experimentation.[6] We use this second inventory of impressions for our experiment, which ensures that ads shown in our experiment are not determined through any algorithmic or human-directed targeting process.

We design a fully randomized experiment at `vdo.ai`, where the pre-roll impressions are assigned to three experimental conditions: (1) No-Ad condition, where the user does not need to watch an ad to start consuming the organic video content, (2) Non-Skippable/Short ad condition, where the ad shown is a 15-second long ad of the `boAt`'s Watch Xtend product that is not skippable, and (3) Skippable/Long ad condition, where the ad shown is a 60-second long ad of the same `boAt` product that is skippable after five seconds. We split the impressions randomly across treatment conditions with different weights, such that the No-Ad condition is used for 10% of all impressions, and either one of the ad conditions is shown in 45% of impressions each. For the Non-Skippable/Short and Skippable/Long ad conditions, we use two versions of an advertisement for the same `boAt` product. The two ads are short and long cuts of the same raw footage, launched by an actual advertiser. In Web Appendix C, we present more details about the advertised product and brand and show a snapshot from each ad format. It is worth emphasizing that although the set of ads is selected from the actual ad inventory, it is not a representative set. As such, our analysis cannot offer generalizable substantive insights about the effect of skippability and length of video ads. Instead, we adopt the perspective of a platform that wants to choose between the two ad formats given its impact on the downstream outcomes, such as sponsored and organic consumption.

Figure 7 summarizes a schema of our experiment and the treatment conditions. The points show the pixels placed to find whether the user has reached a certain point in the ad and video. This means that we can record whether the user has reached the midpoint of the Skippable/Long ad (i.e., second 30) or the third quarter of the organic video (75%). Although we can control for the exact ad shown, the organic

[6]This inventory is unsold only for research and development purposes and does not include impressions that are unsold in the auction. Therefore, the distribution of impressions in this inventory is the same as the impressions that are auctioned off.

Figure 7. A Visual Schema of the Experiment Design.



Note: The experiment is run at the impression level. Step 1 refers to the case where a user generates an impression and the randomization occurs in Step 2.

video content is chosen by the users, so they can be different videos with different lengths. However, given the randomization in our experiment, the organic video does not affect the treatment condition. Thus, the distribution of videos is the same across treatment conditions. We ran the experiment for four days, from July 19–22 in 2022 on a total of 59,692 impressions.

4.2 Data

Each observation in our data refers to a unique session, which is defined as an event where a user visits a web page with video content placed through the `vdo.ai` platform. For each session in our data, we observe a set of *pre-treatment variables*, *treatment assignment*, and a rich set of *post-treatment variables* or *outcomes* as the main focus of our empirical application is to study policy-making when there are many outcomes. For pre-treatment variables, we observe the user's *IP Address*, *Time of Day*, *Date*, *City*, *Country*, and the *Operating System (OS)* of the device that the user is using. For post-treatment variables or outcomes, collect a rich set of post-treatment variables or outcomes both on the ad performance and video engagement metrics. As shown in Figure 7, we place pixels at different points in the sessions that indicate whether the user has reached those points. For both ads, the pixels are placed every 15 seconds[7], and for the video, these pixels are placed for every quarter (25%) of the video.[8] In addition to the pixels shown in Figure 7, we also

[7]It is worth noting that the Ad Consumption in the Non-Skippable/Short condition is recorded at the quarter level. However, we do not use that information to balance the unit of our Ad Consumption outcome across treatments. Our results are robust when we incorporate this information.

[8]It is worth noting that videos can be of different lengths. For example, the first quartile for a two-minute video is reached after 30 seconds, whereas this point can be reached after 10 seconds in a 40-second video. This is a limitation of our analysis. However, the video lengths would not significantly differ across groups since we randomize the treatment.

collect information about whether the user has clicked on the website link embedded in the ad. However, the focal ad in our experiment is a brand ad without a clear click objective. As a result, the click-through rate (CTR) is relatively low.[9] Overall, our rich-feedback environment allows us to evaluate the performance of our treatment conditions in terms of different ad- and video-related metrics used in this industry.

Such detailed tracking also helps us with data cleaning. Specifically, we identify whether a user faces technical issues or uses an ad blocker. In particular, if the pixel at the beginning of both the ad and video returns null values, we assume that the user had technical issues, such as a network problem. Similarly, if the pixel has a null value at the beginning of the ad but a real value at the beginning of the video, we conclude that the user uses ad blockers. Although this approach identifies ad blockers for impressions that are assigned to an ad condition, it is clear that we cannot identify ad blockers for users in the No-Ad condition. However, since our main analysis concerns the difference between two ad formats, this does not cause a problem in our main analysis. Overall, we remove 703 observations because of technical issues and 528 observations for using ad blockers. This gives us a sample of 58,461 observations generated by 56,662 unique users.

For our main analysis, we only focus on the two ad conditions and drop observations for the No-Ad condition, which reduces our sample size to 53,176 sessions generated by 51,423 unique users. Since we only use the first session for each user, our final sample has a total of 51,423 sessions to study. We use this sample throughout the paper for all the results. We present some basic summary statistics of the data. We start with the pre-treatment variables, which are all categorical variables. We find the top three subcategories with the highest number of observations for each variable in our data. We present this information about each variable along with the total number of subcategories in Table 1. As shown in this table, the hours with the highest traffic are 6–8 AM MST, which would be 5:30-7:30 PM in India, where most of the traffic comes from. The experiment was run from July 19 through July 22, and the last two days had the highest traffic.

As indicated in Table 1, there are a total of 956 cities in our data. However, over half of the observations are from Mumbai. It is worth noting that there are many cities with only one observation in our data. Next, we find that the vast majority (99.56%) of all observations occur in India. The statistics on our final

[9] In general, industry reports indicate that the primary focus of digital video ads is to increase brand awareness, as opposed to improving objective performance measures such as click or purchase (Ferguson 2023).

Table 1. Summary Statistics of the Pre-Treatment Variables.

Variable	Number of subcategories	Top three subcategories and their shares		
		1^{st}	2^{nd}	3^{rd}
Hour of Day	24	7AM MST (7.76%)	6AM MST (6.82%)	8AM MST (6.76%)
Date	4	07/21/2022 (39.19%)	07/22/2022 (37.20%)	07/20/2022 (14.69%)
City	956	Mumbai (51.87%)	Delhi (7.55%)	Hyderabad (6.83%)
Country	12	India (99.92%)	United States (0.06%)	Australia (0.00%)
Operating System	6	Android (79.52%)	Windows (13.21%)	iPhone (5.31%)

pre-treatment variable show that Android OS is the most common OS in our data, with around 80% of the total traffic. In Web Appendix D, we perform extensive randomization checks on the distribution of pre-treatment variables to ensure that the randomization has been implemented correctly in our study.

4.3 Average Treatment Effect Analysis

Before we proceed with the analysis, we tie our notation to the methodological framework. We let i denote each observation in our data, and X and W denote the pre-treatment covariates and the treatment variable, respectively. Since we are interested in the difference between the ad formats, we define W as a binary variable with $W = 1$ and $W = 0$ referring to Skippable/Long and Non-Skippable/Short ad formats, respectively.[10] For inference, we use the common assumptions in the causal inference literature: (1) overlap, (2) unconfoundedness, and (3) the stable unit treatment value assumption (SUTVA). The first two are satisfied by design since we have a randomized controlled trial. SUTVA is also reasonable because there is no interaction between users, and the treatment received by all users in the same treatment condition is identical (i.e., no multiple versions of the treatment). Under these assumptions, we know that the average treatment effect is the difference in group averages (Neyman 1923). We use this fact for our main analysis. Here in the main text, we present a summary of average treatment effects on different managerially relevant outcomes as they relate to our multi-objective personalization framework and present the detailed description of outcomes and interpretation of the results in Appendix E.1.

We estimate the average treatment effects on the complete set of outcomes and present the results in Table 2, where each row presents the results for one outcome. The first panel in Table 2 presents the ATE for ad-related outcomes. We find that the average consumption of the Skippable/Long ad is significantly higher than that of the Non-Skippable/Short ad, with the average treatment effect being $0.90 \times 15 = 13.50$ seconds, which is approximately equal to the length of the short ad in our study. We then focus on two other

[10] We only use these two treatment groups in the main text, but we present some results with the No-Ad condition in Web Appendix E.

Table 2. Average Treatment Effects (ATE) Across Outcomes.

Outcome	**Mean of Treatment** A **(Non-Skippable/Short)**	**Mean of Treatment** B **(Skippable/Long)**	**Mean Difference** $B - A$ **Estimate**	p**-value**
Ad Consumption ($\times 15s$)	0.53526	1.43567	0.90041	< 0.001
Second 15 Complete	0.53526	0.55413	0.01887	< 0.001
Ad Complete	0.53526	0.21338	-0.32188	< 0.001
Ad Click	0.00125	0.00128	0.00003	0.926
Video Consumption	0.79714	0.41817	-0.37897	< 0.001
Video Start	0.44535	0.20727	-0.23808	< 0.001
Video Q1 Reached	0.30804	0.15059	-0.15745	< 0.001
Video Q2 Reached	0.22146	0.11634	-0.10512	< 0.001
Video Q3 Reached	0.15768	0.08713	-0.07055	< 0.001
Video Q4 Reached	0.10995	0.06411	-0.04585	< 0.001

Note: The number of observations is 51,423 for all models.

sponsored consumption outcomes: (1) Second 15 Complete and (2) Ad Complete. The former outcome is of interest to platforms that charge Skippable/Long ads once the user watches 15 seconds. Interestingly, we find that the Skippable/Long version has a significantly higher completion of the first 15 seconds than the Non-Skippable/Short version, despite the skippability of the ad in this condition. When we focus on the Ad Complete outcome, we find that the Non-Skippable/Short ad has a substantially higher completion rate, which is expected due to the shorter length and non-skippability of this ad version. The final ad-related outcome in our study is *Ad Click*, which measures whether the user clicked on the ad. We find no significant difference between the Click-Through Rate (CTR) of ad formats, which is expected because the objective of the `boAt` ad campaign in our study is to generate more awareness, not performance measures like clicks.

Next, we focus on video-related outcomes that measure organic consumption. Our main measure of organic consumption is the number of quarters the user has watched the organic content, which we define as *Video Consumption*. We find that users in the Skippable/Long condition consume 0.38 quarters less than those in the Non-Skippable/Short condition. This is equivalent to a 9.5 percentage point difference in the video consumed. We break down Video Consumption into five binary variables corresponding to the beginning of the video and each quarter of the video that is reached. As shown in Table 2, all quarters are more likely to be reached in the Non-Skippable/Short condition than the Skippable/Long condition.

The conflict in average treatment effects on Ad Consumption and Video Consumption raises the question of substitution between the sponsored and organic channels. In order to measure the substitution, one approach is to regress Video Consumption on Ad Consumption to see how the two outcomes are linked.

Table 3. Regression Result for the Link Between Video Consumption and Ad Consumption

	Outcome: Video Consumption (Quarters)	
	(1) Ordinary Least Squares (OLS)	(2) Instrumental Variable (2SLS)
Ad Consumption ($\times$ 15 Seconds)	0.3538^{***}	-0.4209^{***}
	(0.0042)	(0.0150)
Instruments	None	Treatment
Weak Instruments Test		7567^{***}
No. of Obs.	51,423	51,423

Note: Standard errors reported in parenthesis. *p<0.05; **p<0.01; ***p<0.001. The Weak Instrument hypothesis is rejected as the F-statistic in the first-stage regression is 7567 (p<0.001).

However, the main issue with this approach is that users can self-select how much they consume an ad, causing well-known selection or endogeneity bias. We need to use an approach that only uses the exogenous variation in Ad Consumption. Our treatment variable provides a fully random exogenous shifter for this purpose. As a result, we can instrument Ad Consumption with our treatment variable and isolate the causal effect of Ad Consumption on Video Consumption. We present the results from both plain and Instrumental Variable regressions in Table 3. Although the results of column 1 show a positive association between Ad Consumption and Video Consumption in the endogenous specification, we find a strong substitution when we account for endogeneity bias using our 2SLS model, as shown in column 2 of Table 3. Specifically, we find that a 15-second increase in Ad Consumption reduces Video Consumption by 0.42 quarters or 10.52 percentage points. While we do not have the information about the exact length of videos, we know that the average is around 2 minutes or 120 seconds. Using a back-of-the-envelope calculation, we find that every 15-second increase in Ad Consumption decreases Video Consumption by $120 \times 0.1052 = 12.62s$, on average, demonstrating a strong substitution pattern between sponsored and organic consumption.

From a platform perspective, this substitution pattern highlights an inherent trade-off in optimizing sponsored and organic consumption. At the aggregate level, strategies that increase Ad Consumption (sponsored) come at the expense of Video Consumption (organic). As such, this conflict between outcomes creates a perfect setting to provide a proof-of-concept for our multi-objective personalization algorithms.

4.4 Heterogeneous Treatment Effect Analysis

So far, we have shown a strong substitution pattern between ad and video consumption, which poses a challenge for the platform that wants to optimize both outcomes simultaneously. In this section, we explore

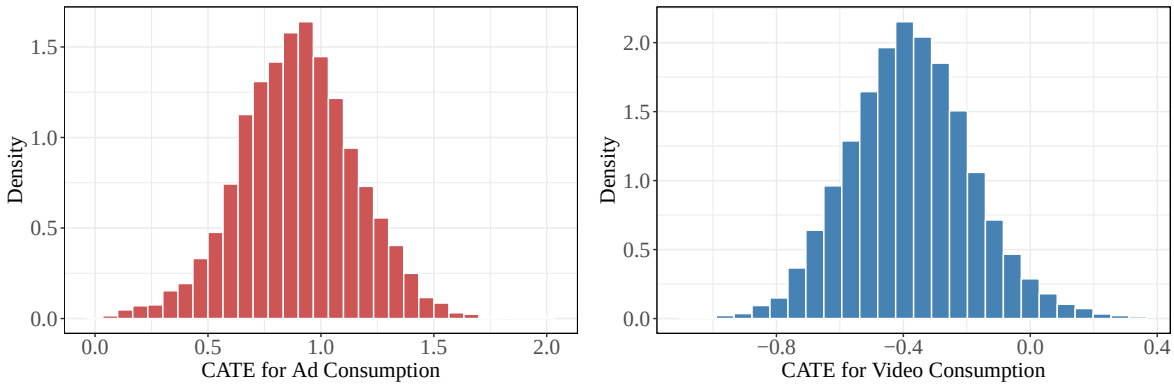
the heterogeneity in treatment effects to see if the substitution pattern persists at a more fine-grained level. As such, we work with the Conditional Average Treatment Effect (CATE) defined in §3.2.1. In principle, if the positive average treatment effect on Ad Consumption and the negative average treatment effect on Video Consumption come from separate portions of our data, the solution is clear for the platform. For example, suppose that there are two groups of users $\mathcal{I}_a$ and $\mathcal{I}_v$ such that $\mathcal{I}_a \cap \mathcal{I}_v = \varnothing$, where users in $\mathcal{I}_a$ have a positive CATE on Ad Consumption and a positive CATE on Video Consumption, whereas users in $\mathcal{I}_v$ have a negative CATE on Ad Consumption and a negative CATE on Video Consumption. In this case, the platform's solution is to assign users $\mathcal{I}_a$ to the Skippable/Long ad and users in $\mathcal{I}_v$ to the Non-Skippable/Short ad. To test this possibility, we need to estimate treatment effects for both outcomes for any individual for a vector of covariates X_i.

In recent years, many methods have been developed to estimate CATE (Shalit et al. 2017, Wager and Athey 2018, Nie and Wager 2021). We use Causal Forests as our main method to estimate CATE on both outcomes. We refer the interested reader to Wager and Athey (2018), Athey et al. (2019) for a detailed algorithm presentation. For the set of covariates, we use all the pre-treatment variables presented in Table 1, as well as the exact timestamp to capture more fine-grained time-dependent heterogeneity and latitude and longitude of cities to go beyond the city categories and capture the spatial heterogeneity patterns (if any). We use 10-fold cross-validation to tune the hyper-parameters of the Causal Forest.

We present the histogram of our CATE estimate for both Ad Consumption and Video Consumption outcomes. Figure 8a shows how CATE for the Ad Consumption outcome varies across individuals. As shown in this figure, although there is extensive variation in the CATE estimates, the sign for all units remains positive. This indicates that the Skippable/Long ad format results in greater Ad Consumption than the Non-Skippable/Short format for all individuals in our data. Thus, if we use this sole objective for developing a personalized policy, the resulting policy will be a uniform Skippable/Long ad condition for everyone.

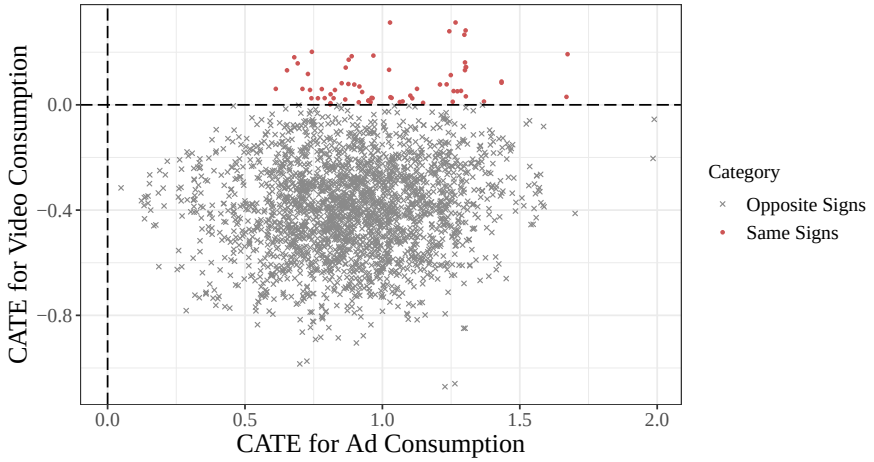
We then move on to CATE estimates for Video Consumption as our video-related outcome and visualize the distribution of CATE estimates in Figure 8b. As shown in this figure, although the vast majority of CATE estimates are negative, there is a small 3.15% of users with positive CATE estimates. Therefore, the optimal personalized policy with respect to Video Consumption as the objective is almost the same as a uniform

Figure 8. The Distribution of CATE Estimates for Ad and Video Consumption.



(a) Outcome: Ad Consumption ($\times 15s$) (b) Outcome: Video Consumption (in Quarters)

Figure 9. Scatter Plot of CATE Estimates on Video Consumption and Ad Consumption



policy where all users are assigned to a Non-Skippable/Short ad. For both outcomes, we present a more interpretable analysis of treatment effect heterogeneity in Web Appendix E.3.

Combining the results from both histograms in Figure 8, although we find substantial variation in the heterogeneous treatment effects on both Ad Consumption and Video Consumption outcomes, the substitution pattern persists even at the individual level. To better understand the substitution pattern at the individual level, we plot the CATE on Video Consumption against the CATE for Ad Consumption and present the resulting scatter plot for a random sample of our observations in Figure 9. The first pattern that emerges from this figure is that only for a small portion of units do we have the same sign for CATE on both outcomes. These points (shown in red in Figure 9) account for 3.15% of all units in our data.

Finally, we ask a broader question: To what extent are CATE estimates for these two outcomes in conflict at the individual level? Since we want higher CATE estimates for both outcomes at the individual level, we

want a more positive correlation between these CATE estimates. On the other hand, a negative correlation between these CATE estimates indicates that a higher CATE for one outcome is associated with a lower CATE for another outcome, thereby making the multi-objective solution more challenging. As shown in Figure 9, there is a weak positive correlation between CATE estimates for both outcomes (correlation = 0.13). Although the positive association between these CATE estimates is not strong, it is still promising as it suggests that CATE estimates move in the same direction, on average. Intuitively, points that contribute most to higher Ad Consumption have a more positive (or less negative) impact on Video Consumption. Thus, the pattern in Figure 9 suggests that multi-objective personalization can be useful for the platform that wants to achieve higher sponsored and organic consumption outcomes simultaneously. To that end, the task at hand is to achieve a good outcome with respect to one objective without compromising too much on the other. We discuss this problem in the next section in greater detail.

5 Returns to Multi-Objective Personalization

5.1 Counterfactual Policy Evaluation

Our main algorithms in §3.2 generate a set of policies. These policies have not been implemented in our data, but we need to evaluate what would have happened had the platform implemented these policies. As such, the question of evaluating a certain policy π becomes one of counterfactual policy evaluation. Since our CATE estimates are structural parameters, we can relatively compare the performance of a policy π with any given baseline policy. Let $\bar{Y}^{(j)}$ denote the average observed for outcome j in the data. We can write:

$$\rho_j(\pi) - \bar{Y}^{(j)} = \frac{1}{N} \sum_{i=1}^{N} \left(\pi(X_i) - W_i\right) \tau_j(X_i), \tag{7}$$

where the elements of this sum are only non-zero when the two policies disagree, i.e., $\pi(X_i) \neq W_i$. Although this approach to policy evaluation has theoretical guarantees such as consistency and unbiasedness, there are a few practical limitations that we must take into account. First, like other high-capacity learners, Causal Forests always face the possibility of overfitting. As a result, we need a reliable approach to evaluate the performance of policies out-of-sample that is robust to overfitting bias. More subtly, even if the CATE estimates do not exhibit overfitting bias, using the same data for policy identification and policy evaluation can result in model-based biases. That is, the policy identifier may exploit the variation in random noise to

generate a policy. If we evaluate the performance of the policy using the same set of estimates, our policy evaluation is subject to the same type of model-based error. Thus, it is important to use a policy evaluation approach that is generalizable and less model-based.

To address this challenge, we use the Inverse Propensity Scoring (IPS) estimator defined in Definition 5 as follows:

$$\hat{\rho}_j^{IPS}(\pi; \mathcal{D}) = \frac{1}{N} \sum_{i \in \mathcal{D}} \left(\frac{\pi(X_i) W_i}{e(X_i)} + \frac{(1 - \pi(X_i))(1 - W_i)}{1 - e(X_i)} \right) Y_i^{(j)}.$$

IPS provides an unbiased estimator of the expected outcome under any policy π. Notably, IPS is a model-free estimator as it does not rely on any outcome model to estimate the outcome under a given policy. Instead, it uses actual outcomes from the data and weights them based on their inverse propensity score to consistently estimate what the expected outcome would have been had policy π been implemented.[11] Another advantage of the IPS estimator is in quantifying the uncertainty. One could use the finite-sample variance of the Horvitz-Thompson estimator to build confidence intervals around the policy evaluation estimates. Since the IPS estimator is defined on the data $\mathcal{D}$, we can easily evaluate both the in-sample and out-of-sample performance of different policies. In particular, we randomly split our data into two sets, where 60% of the observations construct the training data $\mathcal{D}_{\text{Train}}$, and the remaining 40% constitute the test data $\mathcal{D}_{\text{Test}}$. We address the model-based error by performing CATE estimation and policy identification on the training data and evaluating its performance on separate held-out test data. Besides its robustness to model-based errors, our approach is useful as it mimics the practice of real-time policy-making, where the platform uses a batch of data to identify the policies and assign policies in real-time (test data). Thus, platforms can readily apply our framework.

5.2 Policy Identification and Benchmarks

In this section, we identify different sets of policies using the training data and evaluate them on both training and test data. To identify policies using only training data, we need to re-estimate CATE for both Ad Consumption and Video Consumption outcomes on the training data. This ensures that the observations on the held-out test set are not used to estimate CATE. Let $\hat{\tau}_A^{\text{Train}}$ and $\hat{\tau}_V^{\text{Train}}$ denote the estimated CATE functions using the training data for Ad Consumption and Video Consumption outcomes, respectively. We use these estimates to identify different sets of policies. We present a short description of these policies as

[11] Please see Rafieian and Yoganarasimhan (2023) for a detailed explanation of the intuition behind this estimator.

follows, and refer the reader to Web Appendix F for greater details:

- *Scalarization with Causal Estimates* (`MO-SCE`): We use $\mathcal{B} = \{(\beta, 1-\beta) \mid \beta = i/500, 0 \leq i \leq 500\}$ as the full set of weights and run Algorithm 1 using the CATE estimates for both outcomes as inputs. The output is a set of policies denoted by $\Pi^{\text{MO-SCE}}$ containing 501 policies.
- *Scalarization with Policy Learning* (`MO-SPL`): We use the same $\mathcal{B}$, but use the policy learning approach to identify policies. We use $\mathcal{D} = \{X_i, W_i, , e(X_i), \hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i), Y_i^{(A)}, Y_i^{(V)}\}$ as the input for `MO-SPL`.[12] Unlike `MO-SCE`, we do not restrict ourselves to linear models. We use XGBoost as our learning algorithm for the classification task in Algorithm 2. The output is a set of policies denoted by $\Pi^{\text{MO-SPL}}$ containing 501 policies.

For policy comparison and benchmarking, we consider the following policies:

- *Single-Objective for Ad Consumption* (`SO-AC`): This is the personalized policy for Ad Consumption, which we can define as $\pi^{\text{SO-AC}}(X_i) = \mathbb{1}(\hat{\tau}_A^{\text{Train}}(X_i) \geq 0)$. This policy is the same as the `MO-SCE` policy when $\beta = 1$ (weight for the Ad Consumption objective).
- *Single-Objective for Video Consumption* (`SO-VC`): This is the personalized policy for Video Consumption, which we can define as $\pi^{\text{SO-VC}}(X_i) = \mathbb{1}(\hat{\tau}_V^{\text{Train}}(X_i) \geq 0)$. This policy is the same as the `MO-SCE` policy when $\beta = 0$.
- *Mixed Strategy Single Objective* (`SO-Mix`): For any $\alpha \in [0, 1]$, we use a mixed-strategy policy that uses $\pi^{\text{SO-AC}}(\cdot)$ with probability α, and $\pi^{\text{SO-VC}}(\cdot)$ with probability $1 - \alpha$. For consistency with our main policies, we use $\alpha \in \{i/500\}_{i=0}^{500}$ to generate the set of policies $\Pi^{\text{SO-Mix}}$.
- *Random Policy* (`Random`): For any $\alpha \in [0, 1]$, we use a mixed-strategy policy that uses the treatment condition (Skippable/Long) with probability α and the control condition (Non-Skippable/Short) with probability $1 - \alpha$. For consistency, we use $\alpha \in \{i/500\}_{i=0}^{500}$ to generate the set Π^{Random}.

Overall, we have four sets of policies ($\Pi^{\text{MO-SCE}}$, $\Pi^{\text{MO-SPL}}$, $\Pi^{\text{SO-Mix}}$, and Π^{Random}), and two policies ($\pi^{\text{SO-AC}}(\cdot)$ and $\pi^{\text{SO-VC}}(\cdot)$). The single-objective uniform policies and the fully random policies are also helpful benchmarks to use. However, as shown earlier, single-objective personalized policies are almost identical to single-objective uniform policies in our empirical example with Ad Consumption and Video Consumption as the objectives. Thus, to avoid clutter, we do not include these benchmarks.

[12] Normally, one could drop the CATE estimates from inputs. However, the model with CATE estimates ensures a better performance, so we include them as inputs.

5.3 Policy Comparison Results

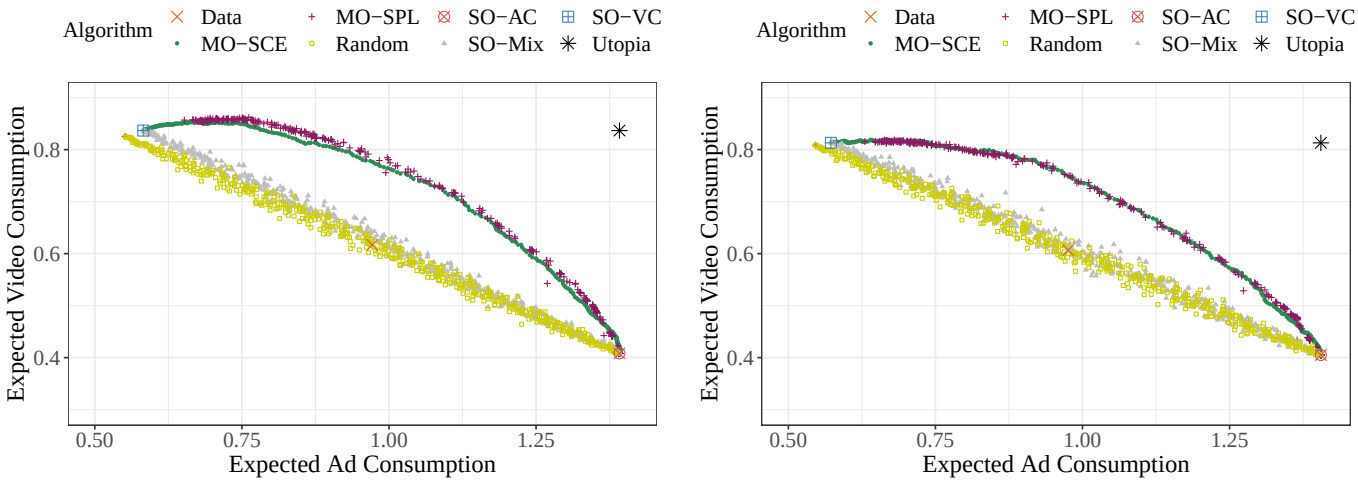
For any policy π, we can use the IPS estimator in Definition 5 to estimate the expected Video Consumption and expected Ad Consumption under that policy on both training and test data. The resulting points are $(\hat{\rho}_A^{IPS}(\pi; \mathcal{D}_{\text{Train}}), \hat{\rho}_V^{IPS}(\pi; \mathcal{D}_{\text{Train}}))$ and $(\hat{\rho}_A^{IPS}(\pi; \mathcal{D}_{\text{Test}}), \hat{\rho}_V^{IPS}(\pi; \mathcal{D}_{\text{Test}}))$ in training and test data, respectively. In addition to the policies described in the previous section, we use two other reference points: (1) Data, which shows the average outcomes for the experiment run in our data, and (2) Utopia, which is the best achievable point and takes the expected outcome under the single-objective personalized policy for that outcome. Figure 10 shows the expected outcomes for these two points, along with the expected outcomes under all policies described in the previous section, separately for the training and test data. Green circles constitute the Pareto frontier under the Scalarization with Causal Estimates (`MO-SCE`) algorithm, whereas purple pluses constitute the Pareto frontier under the Scalarization with Policy Learning (`MO-SPL`) algorithm.[13] Since both single-objective personalized policies are almost the same as the uniform policy, the set of `SO-Mix` and `Random` policies perform very similarly. This is not generally the case if the single-objective personalized policies are non-uniform.

When comparing these Pareto frontiers to the single-objective policies, we see substantial gains as most Mixed Strategy Single Objective (`SO-Mix`) policies are dominated by the Pareto frontier with a relatively large margin. To quantify these gains, we employ two approaches. First, we use the *Covered Area Proportion (CAP)* measure defined in §3.3 and calculate the proportion of the area covered between the Mixed Strategy Single Objective (`SO-Mix`) and Utopia. On the training set, we find 43% and 48% *CAP* measure for `MO-SCE` and `MO-SPL`, respectively. The *CAP* measure is 43% for both policies on the test set. We see a train-test performance discrepancy for `MO-SPL` because it uses a more flexible learner, which is more likely to overfit on the training data, whereas `MO-SCE` uses a linear classification that is less prone to overfitting. The identical performance on the test set is expected as both algorithms optimize the same objective in different ways. Overall, a 43% *CAP* measure on the test set means that the multi-objective personalization algorithms push the Pareto frontier and generate a number of policies that create significant gains in one outcome without sacrificing the other outcome.

Our second approach to demonstrate the value of multi-objective personalization focuses on identifying

[13] The existence of some dominated points within the identified Pareto frontier is expected because we use a separate model-free approach for policy evaluation that captures the randomness in the data.

Figure 10. Policy Evaluation on Training and Test Data



(a) Training Data for Evaluation

(b) Test Data for Evaluation

Note: Each point refers to a certain policy and each policy set is shown with a color and shape. The axes represent the IPS estimates for each outcome under a policy. The confidence intervals are excluded from the figure to the ease of exposition.

policies that improve one outcome without compromising the other one. It is important to notice that a platform cannot simultaneously achieve all the points on the Pareto frontiers shown in Figure 10. This is because the platform can select only one policy. The value of multi-objective personalization is in providing a complete picture for a manager to choose one of the policies on the Pareto frontier that best achieves their objectives. To that end, we highlight three findings that substantially improve one outcome without sacrificing the other:

High Video Consumption, Medium Ad Consumption: From Figure 10a, we see that the manager can choose a variety of policies with great Video Consumption performance while improving expected Ad Consumption. For example, the manager can choose one of the `MO−SCE` policies with $\beta = 0.246$ that results in 4.8% lower Video Consumption compared to the Single-Objective Video Consumption (`SO−VC`) policy, while increasing Ad Consumption by 60.0% on the training data. When comparing the performance of this policy ($\pi^{\text{MO-SCE}}_{0.246}$) with that of the Single-Objective Video Consumption ($\pi^{\text{SO-VC}}$) on the test data, we find that it will result in a drop of 4.4% in Video Consumption while increasing Ad Consumption by 61.0% (from 0.57 to 0.92, or alternatively from 8.58 to 13.82 seconds).

High Ad Consumption, Medium Video Consumption: On the right end of Figure 10a, the manager can choose a policy from `MO−SPL` with $\beta = 0.420$ that achieves a 58.2% improvement in the expected Video Consumption compared to the Single-Objective Ad Consumption (`SO−AC`) policy, while only losing 13.3%

in the expected Ad Consumption. On the test data, the policy $\pi^{\text{MO-SPL}}_{0.420}$ performs 53.2% better in terms of Video Consumption than the ($\pi^{\text{SO-AC}}$) policy, at the expense of 15.2% worse performance in terms of the expected Ad Consumption. Notably, one could consider Video Q1 Reached as a hypothetical point where the platform can place a mid-roll ad. We use the IPS estimator for this outcome to evaluate the proportion of users who reach that point and generate a mid-roll impression under both $\pi^{\text{MO-SPL}}_{0.420}$ and `SO-AC`. Interestingly, we find that the proportion of users who reach the first quarter of the video is 22.7% under $\pi^{\text{MO-SPL}}_{0.420}$. In contrast, this proportion is 14.7% under `SO-AC`, suggesting a 55.5% increase in the number of potential mid-roll ad impressions that would be generated under $\pi^{\text{MO-SPL}}_{0.420}$ relative to `SO-AC`.

High Video Consumption, 15 Seconds Ad Consumption: A useful feature of multi-objective personalization is that we can fix a value for one objective and examine the performance in terms of the other objective. Since Ad Consumption cannot technically be more than 15 seconds in the Non-Skippable/Short ad condition, setting Ad Consumption to 15 seconds would be a reasonable objective. We find that the `MO-SCE` policy with $\beta = 0.274$ achieves 15 seconds of expected Ad Consumption on both training and test sets. We compare the performance of policy $\pi^{\text{MO-SCE}}_{0.274}$ with the two single-objective policies. Compared to the Single-Objective Video Consumption (SOVC) policy, it improves the expected Ad Consumption by 74.0%, while only reducing the Video Consumption by 9.5%, as measured on the test data. On the other end, policy $\pi^{\text{MO-SCE}}_{0.274}$ improves Video Consumption by 81.9% compared to the Single-Objective Ad Consumption (SOAC) policy, while losing 29.2% in Ad Consumption.

In summary, we find that multi-objective personalization results in substantial gains in one objective without sacrificing too much in the other objective. Intuitively, multi-objective personalized policies achieve this by correctly identifying the points in the data whose gains in one objective outweigh their loss in the other objective. From a practical standpoint, platforms can use a batch of data to estimate the primitives, identify the Pareto frontier, and then decide which policy on the Pareto frontier is more desirable.

5.4 Other Case Studies

We demonstrate that the platform can create substantial value by using multi-objective personalization, even in a setting with an almost perfect substitution between the two objectives. As shown earlier, for over 96% of the data points in our data, we observe some degree of substitution between Ad Consumption and Video Consumption. The gains can be significantly larger when the two objectives are less in conflict with each

other. To demonstrate this point, we focus on another set of objectives in Web Appendix G: (1) Second 15 Complete, and (2) Video Consumption. Since many platforms charge advertisers once their ad is watched for 15 seconds (e.g., Facebook), using these two objectives for multi-objective personalization is reasonable for profit-maximizing platforms. We present the results of this practice in Web Appendix G and document substantial gains from a multi-objective personalization policy.

5.5 Implications

Advancements in marketing measurement allow firms to track multiple outcomes of interest at the individual level. For example, a social media platform can measure how much time a user spends on each post, how much ad revenue the user generates, and how many posts the user writes and shares, all of which are presumably outcomes the platform wants to optimize. Naturally, these outcomes have some conflict with each other. We present a general framework for multi-objective personalization that can be used by any manager or decision-maker deciding how to assign interventions to individuals. Our proposed approach empirically identifies the Pareto frontier of policies in terms of multiple outcomes, which helps managers eliminate all the dominated policies. Further, platforms can identify the Pareto optimal policies on a batch of training data and use it as a decision-making dashboard. This feature of our framework allows managers to select the Pareto optimal policy that best balances their objective and then roll it out on the test data.

More specific to our empirical setting, our paper has several implications for video advertising platforms. These platforms often have multiple ad- and video-related objectives, some of which are in direct conflict with the other ones. In our study, we demonstrated a substitution pattern between sponsored and organic content consumption and showed that the platform could create value using our multi-objective personalization framework. Importantly, our framework is fairly general, and a platform can use it across a range of scenarios. For example, if the price per second for Ad Consumption is different across ad formats, the platform needs to rely on a weighted Ad Consumption measure that reflects ad revenues. Our framework can easily accommodate such modifications in the outcome of interest. More broadly, some platforms may be interested in increasing the rate at which the user reaches a certain point within the ad because they charge advertisers based on that rule. In our study, we can consider the 15-second threshold and perform multi-objective personalization for *Second 15 Complete* and *Video Consumption* as our main outcomes of interest (please see Web Appendix G for the results from this practice). More generally, the platform can

have more than two objectives. For example, many streaming platforms also have a subscription-based ad-free version as an alternate revenue channel. As a result, they may be interested in optimizing not only ad and video consumption but also subscription revenue. Our framework can easily be extended to those settings.

Besides offering a prescriptive solution to platforms given the set of objectives, our paper has important market design implications for video advertising platforms. These platforms generally sell ads through auctions. Any auction is characterized by an allocation rule and a payment rule. Our paper highlights why the allocation rule should not be only based on the ad performance but also on the externality it imposes on the system. Prior literature on advertising auctions has studied different forms of ad allocation that capture the externality an ad exposure imposes on other ads (Wilbur et al. 2013, Rafieian 2020). Our paper also suggests another form of externality imposed by ads on content creators, which can affect the supply of ad impressions for the platform in the long run. Platforms can incorporate all these externalities in their allocation and present exact or approximate solutions to this allocation problem. An example of an approximate solution would be capturing these externalities in the quality scores assigned to ads.

These externalities have immediate implications for the payment mechanism in video advertising auctions. In particular, if the platform incorporates the externalities in ad allocation, they need to adjust payments to achieve properties such as truth-telling. Another important implication of our work is for the payment rule in these problems. That is, the platform needs to decide when to charge the advertisers. Some platforms use cutoff-based rules where the advertiser is charged for skippable ads if the user reaches the Second 30 of the ad. Part of the reason for having these rules in place is to account for the externalities an ad exposure can impose on content creators. Given the substitution between ad and video consumption, our findings suggest that a consumption-based payment rule can better account for these externalities. Furthermore, designing an auction with clearer guarantees under a consumption-based payment rule would be easier than in environments with arbitrary cutoff-based rules.

6 Conclusions

Platforms often want to optimize multiple outcomes. A few examples include an online publisher who wants higher sponsored and organic content consumption, a game designer who wants players to play more and pay more, and a social media platform that wants more time spent on the platform, ad revenues, and

user-generated content. Although optimizing all these outcomes seems desirable, finding an optimal intervention for all desired outcomes is often challenging. In some cases, multiple outcomes of interest are in some form of structural conflict. In this paper, we offer personalization as a solution to this problem and propose a multi-objective personalization framework that can reliably identify the Pareto frontier of personalized policies in terms of multiple outcomes. In particular, we propose two algorithms *Scalarization with Causal Estimates* (`MO-SCE`) and *Scalarization with Policy Learning* (`MO-SPL`) that combine the insights from the causal machine learning literature with that of multi-objective optimization and offer solutions with theoretical guarantees. Intuitively, these algorithms exploit the magnitude of the substitution at the individual level to assign individuals to policies.

We apply our framework to a canonical conflict in outcomes between sponsored vs. organic content consumption. In collaboration with `vdo.ai`, we conducted randomized experiments where users were randomly assigned to either the Skippable/Long or Non-Skippable/Short versions of the same ad. We document a high degree of substitution between two key outcomes *Ad Consumption* and *Video Consumption*, even at the individual level. We then apply our multi-objective personalization algorithms and find that the resulting policies improve the outcome in one dimension compared to single-objective personalized policies without sacrificing the outcome in the other dimension. In particular, we show that compared to a single-objective personalized policy that only optimizes Video Consumption, there is a policy on the identified Pareto frontier that improves Ad Consumption by 61.0% while only reducing Video Consumption by 4.4%. Likewise, we document that compared to the single-objective personalized policy that only optimizes Ad Consumption, there is a multi-objective personalized policy that increases Video Consumption by 53.2% while only decreasing the Ad Consumption outcome by 15.2%. We discuss the implications and how the platform can use our framework for optimal decision-making in real-time.

Our paper makes several contributions to the literature. Methodologically, we combine insights from the multi-objective optimization literature with causal machine learning and present a framework for multi-objective personalization. Our proposed algorithms take approaches for Conditional Average Treatment Effect estimation and policy learning and identify a set of policies that are Pareto optimal. From a managerial standpoint, our framework is applicable to other settings where conflicting treatment effects on multiple outcomes are of managerial concern. Our framework offers managers and decision-makers the

flexibility to assess the Pareto frontier separately on a batch of training data and select policies that align with their desired balance between outcomes to roll out. Substantively, we showcase a new area where personalization can create value. Unlike prior literature on personalization that primarily emphasizes the value of personalization through covariate richness, we show how a decision-maker could use the variation in multiple outcomes to differentiate between users and create value through personalization in scenarios where multiple outcomes exhibit substantial conflict. Particularly, we establish notable gains in one outcome without compromising another in a context marked by significant substitution effects. Finally, our paper contributes to the literature on the interplay between sponsored and organic consumption by demonstrating how much personalization can help manage a structural conflict between outcomes. Our findings have important market design implications for advertising platforms, as they highlight the importance of a multi-objective approach for ad allocation and payment rules.

Nevertheless, our paper has limitations that serve as excellent avenues for future research. First, although our proposed scalarization algorithms are scalable for most practical settings where we want to optimize two or three objectives, the set of weights becomes exponentially large as we want to incorporate more objectives. Future research can investigate more efficient ways of using weights as well as approaches to incorporate some domain knowledge to improve the scalability of the algorithm. Second, as is common in the personalization literature, our algorithms require the joint distribution of covariates and potential outcomes to be fixed, abstracting from cases where users act strategically to receive the personalized treatments they prefer in the long run. Future research can extend our framework to such strategic settings similar to Munro (2024). Third, although we use a rich-feedback environment on the logged consumption of ads and videos, we do not have data on whether users pay attention to the screen as in McGranaghan et al. (2022). Using attention data can further illuminate mechanisms behind users' ad and video consumption. Finally, our algorithms involve uncertainty and estimation error. Future research can theoretically examine the generalization error bounds for our algorithms and offer prescriptive solutions for algorithm selection.

Disclosure Statement

The authors certify that they has no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

References

Abdolmaleki, A., Huang, S., Hasenclever, L., Neunert, M., Song, F., Zambelli, M., Martins, M., Heess, N., Hadsell, R., and Riedmiller, M. (2020). A distributional view on multi-objective policy optimization. In *International Conference on Machine Learning*, pages 11–22. PMLR.

Anderson, S. P. and Coate, S. (2005). Market provision of broadcasting: A welfare analysis. *The review of Economic studies*, 72(4):947–972.

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

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

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

Belloni, A., Lovett, M. J., Boulding, W., and Staelin, R. (2012). Optimal admission and scholarship decisions: Choosing customized marketing offers to attract a desirable mix of customers. *Marketing Science*, 31(4):621–636.

Bronnenberg, B. J., Dubé, J.-P., and Mela, C. F. (2010). Do digital video recorders influence sales? *Journal of Marketing Research*, 47(6):998–1010.

Danaher, P. J. (1995). What happens to television ratings during commercial breaks? *Journal of advertising Research*, 35(1):37–37.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197.

Deng, Y. and Mela, C. F. (2018). Tv viewing and advertising targeting. *Journal of Marketing Research*, 55(1):99–118.

Drugan, M. M. and Nowe, A. (2013). Designing multi-objective multi-armed bandits algorithms: A study. In *The 2013 international joint conference on neural networks (IJCNN)*, pages 1–8. IEEE.

Dukes, A., Liu, Q., and Shuai, J. (2022). Skippable ads: Interactive advertising on digital media platforms. *Marketing Science*, 41(3):528–547.

Ferguson, S. (2023). 8 Ways to Increase Brand Awareness With Video in 2023 .

Goli, A., Reiley, D. G., and Zhang, H. (2022). Personalized versioning: Product strategies constructed from experiments on pandora. Working Paper.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685.

Lambrecht, A. and Tucker, C. (2019). Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science*, 65(7):2966–2981.

Marler, R. T. and Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26:369–395.

McGranaghan, M., Liaukonyte, J., and Wilbur, K. C. (2022). How viewer tuning, presence, and attention respond to ad content and predict brand search lift. *Marketing Science*.

Miettinen, K. and Mäkelä, M. M. (2002). On scalarizing functions in multiobjective optimization. *OR spectrum*, 24:193–213.

Munro, E. (2024). Treatment allocation with strategic agents. *Management Science*.

Neyman, J. (1923). Sur les applications de la théorie des probabilités aux experiences agricoles: Essai des principes. *Roczniki Nauk Rolniczych*, 10:1–51.

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

Rafieian, O. (2020). Revenue-optimal dynamic auctions for adaptive ad sequencing. Technical report, Working paper.

Rafieian, O. (2023). Optimizing user engagement through adaptive ad sequencing. *Marketing Science*, 42(5):910–933.

Rafieian, O. and Yoganarasimhan, H. (2021). Targeting and privacy in mobile advertising. *Marketing Science*.

Rafieian, O. and Yoganarasimhan, H. (2023). Ai and personalization. *Artificial Intelligence in Marketing*, pages 77–102.

Rajaram, P., Manchanda, P., and Schwartz, E. M. (2023). Finding the sweet spot: Ad delivery on streaming media. *Available at SSRN 3496039*.

Roijers, D. M., Vamplew, P., Whiteson, S., and Dazeley, R. (2013). A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research*, 48:67–113.

Schweidel, D. A. and Moe, W. W. (2016). Binge watching and advertising. *Journal of Marketing*, 80(5):1–19.

Shalit, U., Johansson, F. D., and Sontag, D. (2017). Estimating individual treatment effect: generalization bounds and algorithms. In *International Conference on Machine Learning*, pages 3076–3085. PMLR.

Siddarth, S. and Chattopadhyay, A. (1998). To zap or not to zap: A study of the determinants of channel switching during commercials. *Marketing Science*, 17(2):124–138.

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

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

Súkeník, P. and Lampert, C. H. (2022). Generalization in multi-objective machine learning. *arXiv preprint arXiv:2208.13499*.

Sun, M. and Zhu, F. (2013). Ad revenue and content commercialization: Evidence from blogs. *Management Science*, 59(10):2314–2331.

Swaminathan, A. and Joachims, T. (2015). Counterfactual risk minimization: Learning from logged bandit feedback. In *International Conference on Machine Learning*, pages 814–823. PMLR.

Tekin, C. and Turğay, E. (2018). Multi-objective contextual multi-armed bandit with a dominant objective. *IEEE Transactions on Signal Processing*, 66(14):3799–3813.

Tuchman, A. E., Nair, H. S., and Gardete, P. M. (2018). Television ad-skipping, consumption complementarities and the consumer demand for advertising. *Quantitative Marketing and Economics*, 16(2):111–174.

Turgay, E., Oner, D., and Tekin, C. (2018). Multi-objective contextual bandit problem with similarity information. In *International Conference on Artificial Intelligence and Statistics*, pages 1673–1681. PMLR.

Van Moffaert, K. and Nowé, A. (2014). Multi-objective reinforcement learning using sets of pareto dominating policies. *The Journal of Machine Learning Research*, 15(1):3483–3512.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 0(0):1–15.

Wang, N., Wang, H., Karimzadehgan, M., Kveton, B., and Boutilier, C. (2022). Imo 3: Interactive multi-objective off-policy optimization. *Proceedings of the 31st International Joint Conference on Artificial Intelligence*.

Wang, Y., Tao, L., and Zhang, X. X. (2023). Recommending for a multi-sided marketplace: A multi-objective hierarchical approach. *Working Paper*.

Wilbur, K. C. (2008). A two-sided, empirical model of television advertising and viewing markets. *Marketing science*, 27(3):356–378.

Wilbur, K. C. (2016). Advertising content and television advertising avoidance. *Journal of Media Economics*, 29(2):51–72.

Wilbur, K. C., Xu, L., and Kempe, D. (2013). Correcting audience externalities in television advertising. *Marketing Science*, 32(6):892–912.

Yoganarasimhan, H., Barzegary, E., and Pani, A. (2022). Design and evaluation of optimal free trials. *Management Science*.

Zufryden, F. S., Pedrick, J. H., and Sankaralingam, A. (1993). Zapping and its impact on brand purchase behavior. *Journal of Advertising Research*, 33(1):58–67.

Web Appendix

A Supplementary Materials for Multi-Objective Personalization Framework

A.1 Proof for Proposition 1

Let $\boldsymbol{\rho}(\pi) \in \mathbb{R}^K$ denote the K-dimensional vector that represents the expected outcomes under policy π, such that $\boldsymbol{\rho}(\pi) = (\rho_1(\pi), \rho_2(\pi), \ldots, \rho_K(\pi))$. Prior literature has shown that if the outcome space is convex, the linear scalarization approach will be able to recover the complete Pareto frontier (Censor 1977, Súkeník and Lampert 2022). Therefore, we only need to show that the outcome space $\{\boldsymbol{\rho}(\pi)\}_{\pi\in\Pi}$ is convex. To do so, we need to show for any two points in the outcome space, all the points on the line between the two also exist in the outcome space. More precisely, for any point $\boldsymbol{r} \in \mathbb{R}^K$ on the line between $\boldsymbol{\rho}(\pi_1)$ and $\boldsymbol{\rho}(\pi_2)$, we need to show that there is a policy π^* such that $\boldsymbol{r} = \boldsymbol{\rho}(\pi^*)$. We know that for any such point $\boldsymbol{r} \in \mathbb{R}^K$ on the line between $\boldsymbol{\rho}(\pi_1)$ and $\boldsymbol{\rho}(\pi_2)$, there exists a $\lambda \in [0, 1]$ such that $\boldsymbol{r} = \lambda\boldsymbol{\rho}(\pi_1) + (1-\lambda)\boldsymbol{\rho}(\pi_2)$. Therefore, it is sufficient to show that we can construct a policy $\pi^* \in \Pi$ that achieves the outcome $\lambda\boldsymbol{\rho}(\pi_1) + (1-\lambda)\boldsymbol{\rho}(\pi_2)$. A natural candidate for π^* is the probabilistic policy that plays a mixed strategy between π_1 and π_2, such that:

$$\pi^*(x) = \lambda\pi_1(x) + (1-\lambda)\pi_2(x) \tag{8}$$

We now show that the expected outcomes under policy π^* is the same as $\boldsymbol{r} = \lambda\boldsymbol{\rho}(\pi_1) + (1-\lambda)\boldsymbol{\rho}(\pi_2)$. Equation (8) shows the probability of treatment assignment under policy $\pi^*(\cdot)$. We can write the probability of control assignment under this policy as follows:

$$1 - \pi^*(x) = \lambda\big(1 - \pi_1(x)\big) + (1-\lambda)\big(1 - \pi_2(x)\big) \tag{9}$$

Now, for any outcome Y_i, we can write:

$$\begin{aligned}
\rho(\pi^*) =& \mathbb{E}\left[\pi^*(X_i)Y_i(1) + \big(1 - \pi^*(X_i)\big)Y_i(0)\right] \\
=& \mathbb{E}\left[\big(\lambda\pi_1(X_i) + (1-\lambda)\pi_2(X_i)\big)Y_i(1) + \Big(\lambda\big(1-\pi_1(X_i)\big) + (1-\lambda)\big(1-\pi_2(X_i)\big)\Big)Y_i(0)\right] \\
=& \mathbb{E}\left[\lambda\Big(\pi_1(X_i)Y_i(1) + \big(1-\pi_1(X_i)\big)Y_i(0)\Big)\right] + \\
& \mathbb{E}\left[\Big((1-\lambda)\Big(\pi_2(X_i)Y_i(1) + \big(1-\pi_2(X_i)\big)Y_i(0)\Big)\Big)\right] \\
=& \lambda\rho(\pi_1) + (1-\lambda)\rho(\pi_2),
\end{aligned} \tag{10}$$

The equation above holds for j's, so we have $\boldsymbol{\rho}(\pi) = \boldsymbol{r}$, and this completes the proof.

A.2 Proof for Proposition 2

Proof. For any vector of covariates x, the joint objective in Equation (2) can be written as follows:

$$\begin{aligned}
\pi_\beta^S &= \operatorname*{argmax}_\pi \sum_{j=1}^K \beta_k \rho_j(\pi) \\
&= \operatorname*{argmax}_\pi \sum_{j=1}^K \beta_j \mathbb{E}\left[\pi(X_i) Y_i^{(j)}(1) + (1-\pi(X_i)) Y_i^{(j)}(0)\right] \\
&= \operatorname*{argmax}_\pi \sum_{j=1}^K \beta_j \mathbb{E}\left[Y_i^{(j)}(0) + \pi(X_i)\left(Y_i^{(j)}(1) - Y_i^{(j)}(0)\right)\right] \\
&= \operatorname*{argmax}_\pi \sum_{j=1}^K \beta_j \left(\rho_j(0) + \tau_j(x)\pi(x)\right) \\
&= \operatorname*{argmax}_\pi \sum_{j=1}^K \beta_j \tau_j(x)\pi(x),
\end{aligned} \tag{11}$$

where the final line uses the fact that $\rho_j(0)$ is policy-invariant. This implies that the optimal policy that maximizes Equation (2) is the one maximizing $\sum_{j=1}^K \beta_j \tau_j(x)\pi(x)$. Therefore, the optimal policy is equal to treatment when $\sum_{j=1}^K \beta_j \tau_j(x) \geq 0$, i.e., $\pi_\beta = \mathbb{1}(\sum_{j=1}^K \beta_j \tau_j(x) \geq 0)$. □

A.3 Proof for Proposition 3

Proof. Since $\hat{\rho}_j^{IPS}(\pi; \mathcal{D})$ is an unbiased estimator for the expected outcome $\rho_j(\pi)$, we can easily see that $\sum_{j=1}^K \beta_j \hat{\rho}_j^{IPS}(\pi; \mathcal{D})$ is also an unbiased estimator for our main objective $\sum_{j=1}^K \beta_j \rho_j(\pi)$ in Equation (2). Therefore, we can write the empirical version of our target $\sum_{j=1}^K \beta_j \rho_j(\pi)$ as follows:

$$\begin{aligned}
\sum_{j=1}^K \beta_j \hat{\rho}_j^{IPS}(\pi; \mathcal{D}) &= \sum_{j=1}^K \beta_j \frac{1}{N_\mathcal{D}} \sum_{i \in \mathcal{D}} \left(\frac{\pi(X_i) W_i}{e(X_i)} + \frac{(1-\pi(X_i))\,(1-W_i)}{1-e(X_i)}\right) Y_i^{(j)} \\
&= \frac{1}{N_\mathcal{D}} \sum_{i \in \mathcal{D}} \left(\frac{\pi(X_i) W_i}{e(X_i)} + \frac{(1-\pi(X_i))\,(1-W_i)}{1-e(X_i)}\right) \left(\sum_{j=1}^K \beta_j Y_i^{(j)}\right)
\end{aligned} \tag{12}$$

The expression above implies that we can think of the optimization problem in Equation (2) as a single-objective policy learning problem with the outcome being $\sum_{j=1}^K \beta_j Y_i^{(j)}$. Therefore, this policy learning problem is equivalent to single-objective policy learning. We now use the derivation in Athey and Wager (2021) to show that the optimal solution to this problem is the same as the optimal solution to a weighted classification problem. For brevity, we drop $\mathcal{D}$ from $\hat{\rho}_j(\cdot; \mathcal{D})$. Let $\rho_j(0)$ and $\rho_j(1)$ denote the expected outcome j under the uniform control and uniform treatment policies, respectively. We note that optimizing $\operatorname{argmax}_\pi \sum_{j=1}^K \beta_j \rho_j(\pi)$ is the same as optimizing its advantage over uniform policies $\operatorname{argmax}_\pi \sum_{j=1}^K 2\beta_j \rho_j(\pi) - \sum_{j=1}^K \beta_j(\rho_j(1) + \rho_j(0))$, since the second summation is policy invariant. There-

fore, we can rewrite our target objective as follows:

$$
\begin{aligned}
\underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \hat{\rho}_j^{IPS}(\pi) &= \underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} 2\beta_j \hat{\rho}_j^{IPS}(\pi) - \sum_{j=1}^{K} \beta_j \left(\hat{\rho}_j^{IPS}(1) + \hat{\rho}_j^{IPS}(0)\right) \\
&= \underset{\pi}{\operatorname{argmax}} \frac{1}{N_{\mathcal{D}}} \sum_{i \in \mathcal{D}} 2 \left(\frac{\pi(X_i) W_i}{e(X_i)} + \frac{(1-\pi(X_i))\,(1-W_i)}{1-e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \\
&\qquad - \frac{1}{N_{\mathcal{D}}} \sum_{i \in \mathcal{D}} \left(\frac{W_i}{e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \\
&\qquad - \frac{1}{N_{\mathcal{D}}} \sum_{i \in \mathcal{D}} \left(\frac{(1-W_i)}{1-e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \\
&= \underset{\pi}{\operatorname{argmax}} \frac{1}{N_{\mathcal{D}}} \sum_{i \in \mathcal{D}} \left(\frac{(2\pi(X_i)-1) W_i}{e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \\
&\qquad + \frac{1}{N_{\mathcal{D}}} \sum_{i \in \mathcal{D}} \left(\frac{(1-2\pi(X_i))\,(1-W_i)}{1-e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \\
&= \underset{\pi}{\operatorname{argmax}} \left(2\pi(X_i)-1\right) \left(\frac{W_i}{e(X_i)} - \frac{(1-W_i)}{1-e(X_i)} \right) \left(\sum_{j=1}^{K} \beta_j Y_i^{(j)} \right) \\
&= \underset{\pi}{\operatorname{argmax}} \left(2\pi(X_i)-1\right) \Gamma_i \\
&= \underset{\pi}{\operatorname{argmax}} \underbrace{\left(2\pi(X_i)-1\right) L_i}_{\text{Classification Objective}} \quad \underbrace{|\Gamma_i|}_{\text{Observation Weight}}
\end{aligned}
\tag{13}
$$

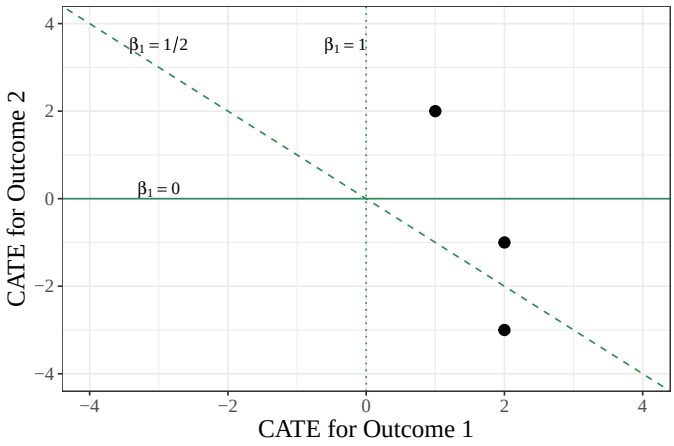
The alternative objective in the first line allows us to use simple algebraic simplifications to reach a classification objective $(2\pi(X_i) - 1)\, L_i$. Therefore, we show that optimizing the scalarized objective is equivalent to the weighted classification in Proposition 3 and complete the proof.

B Illustrative Examples for Algorithms

B.1 Illustrative Example for the Scalarization with Causal Estimates Algorithm

Consider a simple case with three data points. For each point, we have CATE estimates for two outcomes. Let T denote the set of CATE estimates on both outcomes for all three points, such that $T = \{(1, 2), (2, -1), (2, -3)\}$. We want to run our Scalarization with Causal Estimates (`MO-SCE`) algorithm. We set $\mathcal{B}$ such that $\beta_1 \in \{0, 1/2, 1\}$. As such, our `MO-SCE` algorithm generates three separate policies. Figure A1 illustrates the lines that determine these three policies. We discuss each below:

Figure A1. Illustrative Example for the Scalarization with Causal Estimates Algorithm



- *Case 1:* $\beta_1 = 0$. In this case, the weight for Outcome 1 is zero, and the weight for Outcome 2 is one. The data point $(1, 2)$ is above the line (i.e., it satisfies $\beta_1\tau_1 + (1 - \beta_1)\tau_2 \geq 0$ for $\beta_1 = 0$), so it will be assigned to the treatment condition. The two other points will be assigned to the control condition.
- *Case 2:* $\beta_1 = 1/2$. In this case, the weight for both outcomes is 1/2. As shown in Figure A1, points $(1, 2)$ and $(2, -1)$ are above the line corresponding to $\beta_1 = 1/2$, so these two points will be assigned to the treatment condition. The point $(2, -3)$ is below the line, so it will be assigned to the control condition.
- *Case 3:* $\beta_1 = 1$. In this case, we only consider Outcome 1. As shown in Figure A1, all three points satisfy $\beta_1\tau_1 + (1 - \beta_1)\tau_2 \geq 0$, so they will all be assigned to the treatment condition.

Overall, the `MO-SCE` algorithm provides three policies as the output. These policies correspond to different points on the Pareto frontier of the outcome space.

B.2 Illustrative Example for the Scalarization with Policy Learning Algorithm

We now provide an illustrative example for the Scalarization with Policy Learning (`MO-SPL`) algorithm. We revisit the same stylized example in Appendix B.1, where $T = \{(1, 2), (2, -1), (2, -3)\}$. The illustrative example in this case is a little more challenging to present as we need other pieces such as the actual treatment assignment in the data for the data points and their outcomes. Suppose the vector of treatment assignments in the data is $W = (1, 0, 1)$, such that points $(1, 2)$ and $(2, -3)$ are assigned to the treatment with the propensity score $1/2$, and the point $(2, -1)$ is assigned to the control condition with the propensity score $1/2$. For outcomes, we consider a simple case where $Y^{(j)} = 1 + W \odot \tau_j$ for Outcome $j \in \{1, 2\}$, where $\odot$ is the element-wise product. As such, we have the following outcomes: $Y^{(1)} = 1 + (1, 0, 1) \odot (1, 2, 2) = (2, 1, 3)$ and $Y^{(2)} = 1 + (1, 0, 1) \odot (2, -1, -3) = (3, 1, -2)$.

To run our `MO-SPL` algorithm, we set $\mathcal{B}$ where $\beta_1 \in \{0, 1/2, 1\}$. For each β_1, we generate a separate policy. We discuss how these policies are generated in Algorithm 2 as follows:

- *Case 1:* $\beta_1 = 0$. In this case, the weight for Outcome 1 is zero, and the weight for Outcome 2 is one. We first calculate Γ in this case, which consists of two parts: (1) a difference between the propensity

weights $W_i/e(X_i) - (1-W_i)/(1-e(X_i))$, and (2) a weighted outcome $\sum_{j=1}^{K} \beta_j Y_i^{(j)}$. For data points assigned to the treatment, the difference between the propensity weights is $1/(0.5) - 0/(0.5) = 2$, and for data points assigned to the control, this difference is $0/(0.5) - (1-0)/(0.5) = -2$.

$$\Gamma = (2 \times 3, -2 \times 1, 2 \times -2) = (6, -2, -4) \tag{14}$$

Therefore, the labels will be $(1, -1, -1)$ with weights $(6, 2, 4)$. We can then run this classification with any set of covariates we want.

- *Case 2:* $\beta_1 = 1/2$. In this case, the weight for both outcomes is 1/2. We can calculate Γ as follows:

$$\Gamma = \left(2 \times \left(\frac{1}{2}2 + \frac{1}{2}3\right), -2 \times \left(\frac{1}{2}1 + \frac{1}{2}1\right), 2 \times \left(\frac{1}{2}3 + \frac{1}{2}(-2)\right)\right) = (5, -4, 1) \tag{15}$$

Therefore, the labels will be $(1, -1, 1)$ and the weights for our classification task are $(5, 4, 1)$.

- *Case 3:* $\beta_1 = 1$. In this case, we only consider Outcome 1. We calculate Γ as follows:

$$\Gamma = (2 \times 2, -2 \times 1, 2 \times 3) = (2, -2, 6) \tag{16}$$

The labels will be $(1, -1, 1)$, similar to the previous case, but the weights for our classification task are $(2, 2, 6)$.

Overall, the `MO-SPL` provides three policies as the output. These policies correspond to different points on the Pareto frontier of the outcome space.

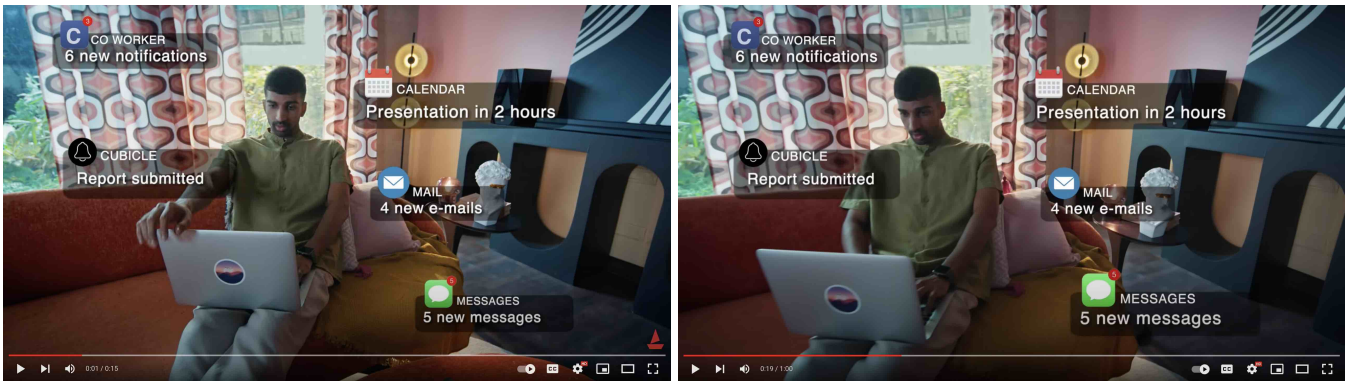
C Description of `boAt` Ads in the Experiment

In this section, we present supplementary information about the advertised brand and product. `boAt` is an audio and wearables brand that offers different wireless earphones, earbuds, headphones, smartwatches, and home audio. The advertised product, Watch Xtend, is a smartwatch with a built-in Alexa voice assistant. The ad content in both conditions presents a series of takes where different characters use the Alexa voice assistant in different contexts, highlighting the utilitarian and lifestyle aspects of the product. The primary objective of the ad campaign is to generate greater awareness about the product and brand. As such, although the ad is clickable, it is not a performance ad with a clear call for action. A click only takes users to the product's website for more information. Figure A2 shows a snapshot of different parts of both ad versions.

D Randomization Check

In this section, we use the pre-treatment variables to check whether randomization in our experiment has been implemented properly. In particular, we want to perform randomization checks on the sample used for our main analysis. To do so, we need to check if there is any discrepancy in the distribution of the pre-treatment covariates across the two treatment conditions we use in our main analysis: Skippable/Long and Non-Skippable/Short ad formats. In our setting, all the pre-treatment variables are categorical. As

Figure A2. Screenshots of Non-Skippable/Short and Skippable/Long Ads Used in the Experiment.



(a) Non-Skippable/Short (b) Skippable/Long

such, we want to examine if there is any significant difference in the proportions of each subcategory under Skippable/Long and Non-Skippable/Short conditions. We use three different approaches to assess randomization checks in our data:

- Hypothesis Testing Approach: Our first approach uses the hypothesis testing framework. For any subcategory s, let $\pi_{s,1}$ and $\pi_{s,0}$ denote the proportion of observations belonging to the subcategory in the Skippable/Long and Non-Skippable/Short conditions, respectively. If randomization has been done correctly, we will fail to reject the following null hypothesis: $H_0 : \pi_{s,0} = \pi_{s,1}$. First, we conduct Fisher's exact test for each subcategory. Since we run multiple hypotheses, we expect a fraction of them to be significant even if the null hypothesis is true. Of 837 separate tests conducted, only 8 rejected the null hypothesis. After adjusting for multiple hypothesis testing using the Benjamini-Hochberg approach (Benjamini and Hochberg 1995), no adjusted p-value was below 0.05.[14] In Table A1, we present the proportions under Skippable/Long and Non-Skippable/Short for the top subcategories identified in Table 1, as well as the Fisher and Z-test p-values for their corresponding hypothesis tests. As shown in this table, all the p-values are greater than 0.05, indicating that we fail to reject the null hypothesis that the proportions are the same under Skippable/Long and Non-Skippable/Short conditions. This evidence suggests that randomization has been implemented correctly in our data.
- Standardized Bias Approach: In our second approach, we use the measure of *Standardized Bias (SB)*, which is commonly used in the literature to assess covariate balance. Standardized Bias is equal to the absolute difference between the means of two groups divided by the standard deviation of the covariate for the pooled sample. The common norm in the literature is to consider a Standardized Bias below 0.2 or 0.1 as evidence for covariate balance (McCaffrey et al. 2013). In our setting, we find that the maximum Standardized Bias was 0.026, which indicates that we have a covariate balance for all the pre-treatment covariates using this approach. In Table A1, we present the Standardized Bias for the top subcategories. As shown in the table, all values are substantially lower than the acceptable thresholds.
- Regression Approach: We use a regression approach to regress the treatment assignment on all the pre-treatment variables. If randomization has been done correctly, the pre-treatment variables will have

[14] We arrive at the same conclusion when we use Z-test for comparing two proportions.

Table A1. Randomization Checks for Top Subcategories

Subcategory	Proportion Under		p-value		Standardized
	NS	SL	Fisher	Z-test	Bias
Hour of Day: 7AM (MST)	0.0776	0.0777	0.9606	0.9578	0.0005
Hour of Day: 6AM (MST)	0.0687	0.0676	0.6242	0.6140	0.0044
Hour of Day: 8AM (MST)	0.0675	0.0677	0.9440	0.9387	0.0007
Date: 07/21/2022	0.3916	0.3921	0.9136	0.9124	0.0010
Date: 07/22/2022	0.3718	0.3722	0.9128	0.9076	0.0010
Date: 07/20/2022	0.1479	0.1459	0.5090	0.5083	0.0058
City: Mumbai	0.5194	0.5180	0.7506	0.7461	0.0029
City: Delhi	0.0738	0.0772	0.1416	0.1409	0.0130
City: Hyderabad	0.0698	0.0668	0.1782	0.1745	0.0120
Country: India	0.9990	0.9993	0.2854	0.2289	0.0106
Country: US	0.0006	0.0006	1.0000	0.9842	0.0002
Country: Australia	0.0001	0.0000	0.5004	0.1573	0.0121
OS: Android	0.7974	0.7929	0.2167	0.2135	0.0110
OS: Windows	0.1307	0.1335	0.3550	0.3522	0.0082
OS: iPhone	0.0533	0.0530	0.9217	0.9123	0.0010

Note: NS and SL are short acronyms for Non-Skippable/Short and Skippable/Long, respectively

no predictive power in explaining the treatment assignment. We can statistically test that by using the F-test of the regression model. We find that the F-statistic is equal to 1.02 with a p-value of 0.32, which indicates that the pre-treatment variables have no predictive power in predicting the treatment assignment and provides evidence for the validity of randomization in our study.

It is worth emphasizing that the randomization checks have all been performed on the focal sample used for the main analysis. However, as discussed in the main text, we use some sampling to arrive at our final data, such as dropping users who use ad blockers or face technical issues. Theoretically, we expect our randomization checks to hold in the original data, as the actual randomization was implemented on that data with the observations that were dropped in our sample. We perform the same analysis for the original sample and arrive at the same insight that randomization has been correctly performed.

E Supplementary Material for Experiment Analysis

E.1 Outcome Variables

We have two sets of outcomes: (1) ad-related outcomes and (2) video-related outcomes. The former demonstrates user behavior regarding the ad (e.g., how much ad content to consume, click), whereas the latter captures user behavior regarding the video (e.g., how much video content to consume). Table A2 presents a list of our outcome variables along with their description. The first four outcome variables are ad-related, whereas outcomes 5–10 are video-related. For each outcome variable Y, we consider a set of potential outcomes $Y(w)$, where w is the value of our treatment variable.

Table A2. Description of Outcome Variables.

No.	Outcome	Description
1	Ad Consumption	Numerical variable measuring how much ad content the user has consumed in discrete 15s units.
2	Second 15 Complete	Binary variable indicating whether the user has reached the 15^{th} second of the ad.
3	Ad Complete	Binary variable indicating whether the user has completed watching the entire ad.
4	Ad Click	Binary variable indicating whether the user has clicked on the ad.
5	Video Consumption	Numerical variable indicating how many quarters of the video have been watched by the user.
6	Video Start	Binary variable indicating whether the user has started watching the organic video content.
7	Video Q1 Reached	Binary variable indicating whether the user has reached the 1st quarter (25%) of the video.
8	Video Q2 Reached	Binary variable indicating whether the user has reached the 2nd quarter (50%) of the video.
9	Video Q3 Reached	Binary variable indicating whether the user has reached the 3rd quarter (75%) of the video.
10	Video Q4 Reached	Binary variable indicating whether the user has reached the 4th quarter (100%) of the video.

E.2 Interpretation of Average Treatment Effect

In this section, we provide a more extensive interpretation of findings presented in Table 2. We expand on our earlier discussion in the main text and provide a more in-depth analysis and interpretation of results. We separate the analysis by outcomes and present the discussion for each separately as follows:

Ad Consumption: Theoretically, it is unclear which ad format leads to higher Ad Consumption. On the one hand, a longer ad has an inherent advantage as it can be consumed for a longer time. On the other hand, the ability to skip the long ad after 5 seconds may result in lower consumption of the longer ad. As indicated in Table 2, we find that the average consumption of the Skippable/Long ad is significantly higher than that of the Non-Skippable/Short ad, with the average treatment effect being $0.90 \times 15 = 13.50$ seconds, which is approximately equal to the length of the short ad in our study.

Second 15 Complete: We use the binary outcome *Second 15 Complete*, as defined in Table A2. With the same length of consumption, we can better examine the role of the skippability option, as users in only one condition can skip the ad. The conventional wisdom is that the Skippable/Long ad format will be less likely to consume 15 seconds of the ad. Surprisingly, we find the opposite pattern in the second row of Table 2: despite the presence of the skip option in the Skippable/Long condition, the completion rate of the first 15 seconds is significantly higher in Skippable/Long condition compared to the Non-Skippable/Short condition. One outcome that explains this effect is the low ad skip rate: we find that only 1.9% of users in the Skippable/Long ad condition skip the ad before the Second 15 checkpoint to start watching the video. This is different from other video streaming contexts, such as YouTube, where users skip ads at a very high rate. However, this could explain a null treatment effect. The significantly higher *Second 15 Complete* rate under Skippable/Long compared to Non-Skippable/Short can be attributed to other factors, such as the difference in the video content for the first 15 seconds.

Ad Complete: In the third row of Table 2, we estimate the treatment effect on the outcome *Ad Complete*. Theoretically, we expect a higher ad completion under Non-Skippable/Short ad because shorter ads are easier to complete, and the inability to skip forces users in this condition to complete the ad in order to watch the organic video content. As expected, we find that 53.5% of Non-Skippable/Short ads are completed, whereas only 21.3% of Skippable/Long ads are completed. The difference is largely significant.

Ad Click: We focus on users' click decision on ads as the final outcome (*Ad Click*). As discussed earlier, the objective of the `boAt` ad campaign in our study is to generate more awareness. As such, although the ad is clickable, it is not a performance ad with a clear call for action. A click only takes users to the product's website for more information. The fourth row of Table 2 compares the performance of the two conditions in terms of Ad Click. Both ads generate around 0.1% click-through rate (CTR), and the difference is not statistically significant.

Video Consumption: Since we work with pre-roll ads, we expect the ad format to affect Video Consumption. As indicated in Table 2, we find that users in the Skippable/Long condition consume 0.38 quarters less than those in the Non-Skippable/Short condition. This is equivalent to a $0.38 \times 0.25 \times 100 = 9.5$ percentage point difference in the video consumed. Our results are different from prior studies on ad skippability have documented a higher rate of organic content consumption when users are exposed to skippable ads (Pashkevich et al. 2012). The reason for the difference likely comes from the fundamental difference between our setting and YouTube, where Pashkevich et al. (2012) conducted their study. In our setting, the organic video is not necessarily the primary reason why a user is in a session. As such, the more time they spend on advertising content, the less they have for the organic video, suggesting a substitution pattern between these outcomes.

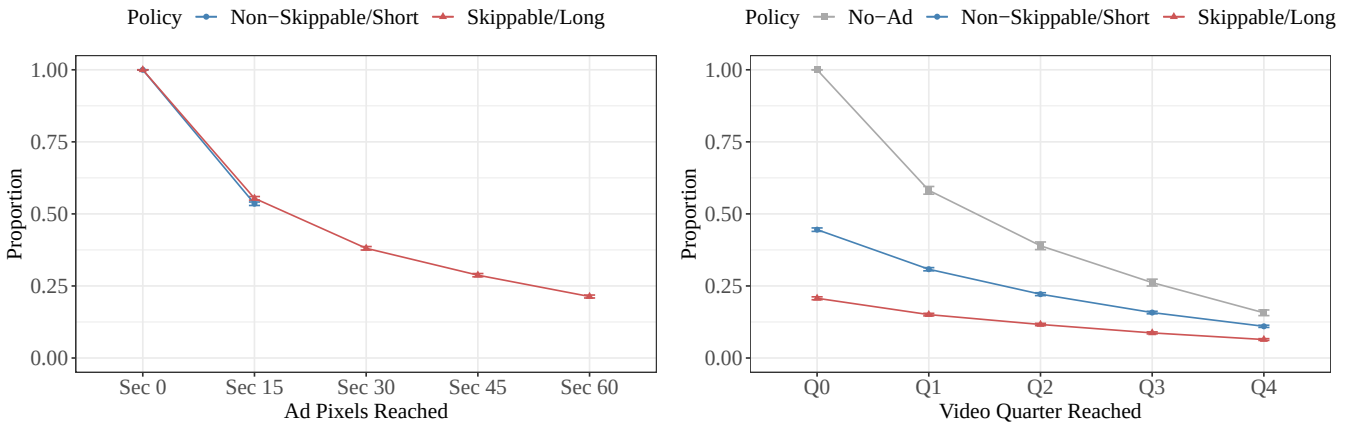
Video Start: We find that users in the Non-Skippable/Short ad condition are more likely to start the video than users in the Skippable/Long ad condition. Because users in the Skippable/Long ad condition spend more time on the ad, they are less likely to be present at the beginning of the organic video. Further, we note that 44.5% of users in the Non-Skippable/Short condition started watching the video, which is lower than 53.5% who completed the ad, indicating that there is some dropout in the transition from ad to video.

Video Qj Reached: For any video quarter j, we define a binary variable Video Qj Reached to break down the Video Consumption variable and measure the treatment effects at different points. We find that the Skippable/Long ad results in a lower Video Qj Reached compared to the Non-Skippable/Short ad for any j. The substitution pattern between Ad Consumption and Video Consumption can explain this pattern.

We now visually summarize all the findings in Figure A3. In Figure A3a, we show the proportion of users who reach a certain pixel for both Skippable/Long and Non-Skippable short ad formats. We notice the higher survival rate at the completion of the Second 15 pixel for the Skippable/Long condition. This figure also shows a decay rate at the subsequent pixels for the Skippable/Long ad condition.

Next, in Figure A3b, we visualize the proportion of users who reach a certain video pixel for Skippable/Long and Non-Skippable/Short ad conditions, as well as these values for the No-Ad condition. Please note that we cannot show the ad-related outcomes for the No-Ad condition, as there is no ad shown by design. As a result, the video start rate is 1, and the fraction of surviving users decreases over the course of the video. The fraction of users who reached each quarter of the video is significantly higher for the control condition than both the conditions with a pre-roll ad before the video. This finding highlights the role of ad avoidance in our study, as both conditions with an ad result in substantially lower Video Consumption. Another interesting pattern that emerges from Figure A3b is the difference in the rates at which the fraction of surviving users declines across different policies. This figure shows that the No-Ad condition has the

Figure A3. Proportion and Confidence Intervals for Binary Outcomes Across Experimental Conditions



(a) Ad-Related Outcomes

(b) Video Related Outcomes

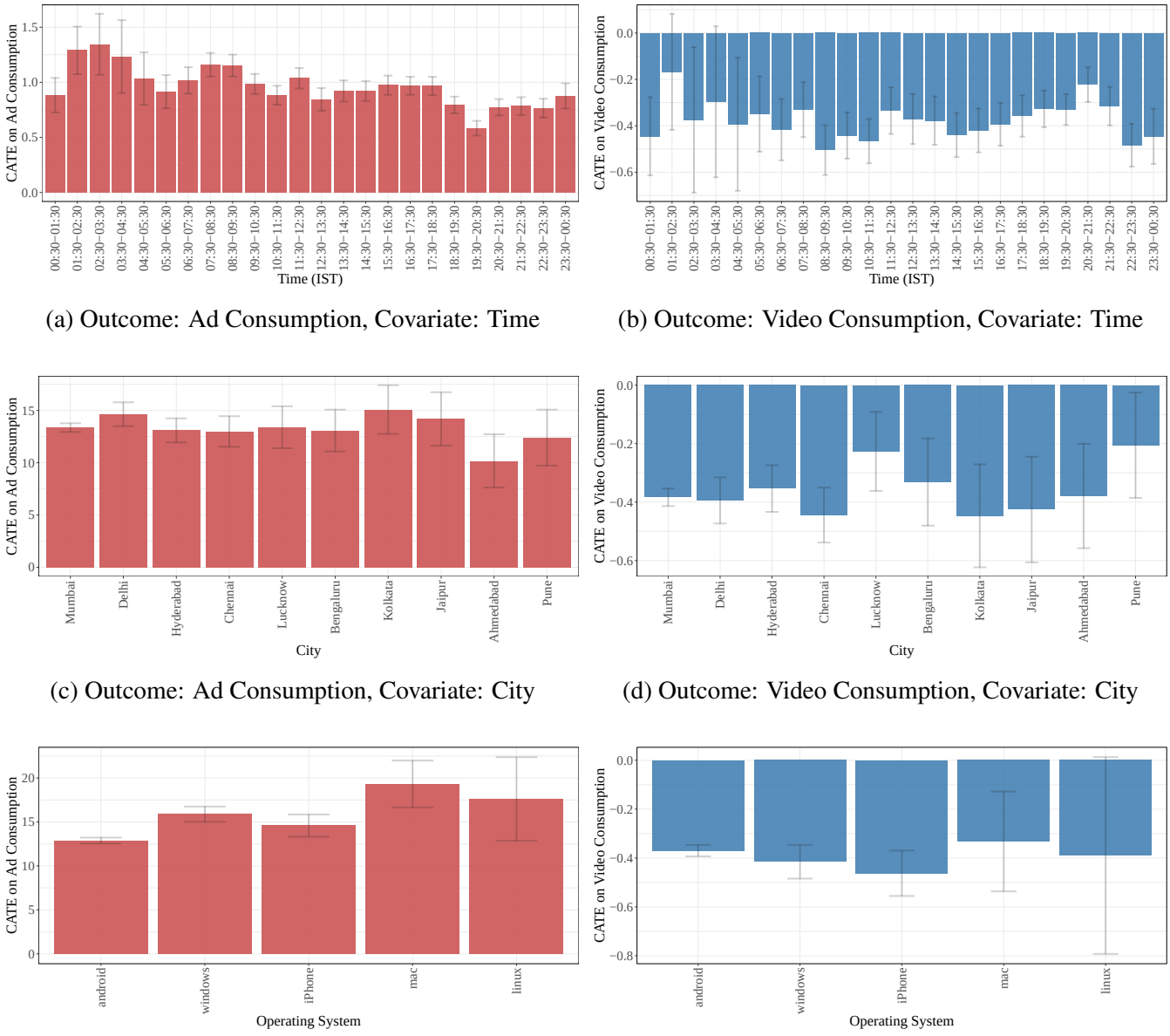
steepest negative slope, whereas the Skippable/Long ad condition has the flattest negative slope. It is worth emphasizing that this is not a causal effect, and there is selection in the type of users who are start the video across conditions.

E.3 Heterogeneity in Treatment Effects Across Pre-Treatment Variables

In this section, we visualize the heterogeneous treatment effects across pre-treatment covariates. In particular, we focus on (1) Hour of Day, (2) City, and (3) Operating Systems, and quantify the treatment effect conditional on different subcategories in these pre-treatment variables. Because we have an experiment, we know that treatments are properly randomized at any given point in time. As a result, we can use a simple mean difference estimator for the data from each subcategory. We only focus on our sample in India as it constitutes 99.5% of all observations, and we modify the time zone from MST to IST for a more meaningful interpretation. We focus on the top ten cities in terms of session count and drop Cros from Operating Systems as there are only 36 observations belonging to users with Cros as the OS.

We present the results in Figure A4. Although we find heterogeneity in treatment effects across subcategories, the substitution pattern persists across all subcategories. We now present some speculative interpretations for the variation observed in Figure A4. The magnitude of treatment effects on both outcomes is lower after 5:30 PM till 8:30 PM. This indicates that users' consumption is overall less sensitive to the ad format. One possible explanation for this pattern is that it is during the more focused leisure time of the users. The extent of substitution is at its lowest during midnight hours and highest during regular work hours, which is reasonable given the difference in time flexibility during midnight relative to work hours. Focusing on the cities, we see some heterogeneity, with some cities having a lower CATE on Video Consumption. By and large, we see a lower substitution level for less industrial cities, which could be due to their time flexibility. Lastly, we see that the treatment effect on Ad Consumption is higher for desktop users (Mac and Windows) compared to mobile users (Android and iPhone), which is likely because mobile sessions are shorter and, therefore, more likely to be abandoned quickly.

Figure A4. Heterogeneity in Treatment Effects on Both Outcomes Across Pre-Treatment Variables



(a) Outcome: Ad Consumption, Covariate: Time

(b) Outcome: Video Consumption, Covariate: Time

(c) Outcome: Ad Consumption, Covariate: City

(d) Outcome: Video Consumption, Covariate: City

(e) Outcome: Ad Consumption, Covariate: OS

(f) Outcome: Ad Consumption, Covariate: OS

Note: Error bars are 95% confidence intervals around treatment effects. Ad Consumption is measured in 15-Second units and Video Consumption is measured in Quarters. In the first row, Times are presented in Indian Standard Time (IST). In the second row, we present top ten cities based on the session count, from left to right.

F Details of Policies Defined in §5.2

In this section, we present a detailed and more formal version of the policies defined in §5.2.

Scalarization with Causal Estimates (`MO-SCE`): We first define $\mathcal{B} = \{(\beta, 1-\beta) \mid \beta = i/500, 0 \leq i \leq 500\}$ as the full set of weights. For any $\beta \in \mathcal{B}$, we need to define the policy $\pi_\beta^{\text{MO-SCE}}$. We use the following equation to define this policy:

$$\pi_\beta^{\text{MO-SCE}}(X_i) = \mathbb{1}\left(\beta\hat{\tau}_A^{\text{Train}}(X_i) + (1-\beta)\hat{\tau}_V^{\text{Train}}(X_i) \geq 0\right), \tag{17}$$

where $\hat{\tau}_A^{\text{Train}}(X_i)$ and $\hat{\tau}_V^{\text{Train}}(X_i)$ are CATE estimates on Ad Consumption and Video Consumption outcomes. We enumerate over all possible values of $\beta \in \mathcal{B}$ and form the set of policies $\Pi^{\text{MO-SCE}}$, which contains 501 policies.

Scalarization with Policy Learning (`MO-SPL`): We use the same $\mathcal{B}$ to choose the weights from. We use $\mathcal{D} = \{X_i, W_i, , e(X_i), \hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i), Y_i^{(A)}, Y_i^{(V)}\}$ for observations in the training data, where $e(X_i)$ is the propensity score for treatment assignment and $Y_i^{(A)}$ and $Y_i^{(V)}$ denote the Ad Consumption and Video Consumption outcomes, respectively. For any specific $\beta \in \mathcal{B}$, we can define $\Gamma_i^{(\beta)}$ as follows:

$$\Gamma_i^{(\beta)} = \left(\frac{W_i}{e(X_i)} - \frac{1-W_i}{1-e(X_i)}\right)\left(\beta Y_i^{(A)} + (1-\beta)Y_i^{(V)}\right) \tag{18}$$

We then define $L_i^{(\beta)} = sgn(\Gamma_i^{(\beta)})$ as the label and $|\Gamma_i^{(\beta)}|$ as the observation weights. The next step is to run a classification algorithm with $L_i^{(\beta)}$ as the binary outcome, $\{X_i, \hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i)\}$ as covariates, and $|\Gamma_i^{(\beta)}|$ as observation weights. We use XGBoost for this classification task, which estimates the function $\hat{f}_{\text{XGB}}^{(\beta)}(\cdot)$ that outputs a binary decision that takes value one if the policy suggests treatment, and zero otherwise. We use a cross-validation procedure to estimate this function on the training set and hold out the test set throughout the process. We can define policy $\pi_\beta^{\text{MO-SPL}}$ as follows:

$$\pi_\beta^{\text{MO-SCE}}(X_i) = \hat{f}_{\text{XGB}}^{(\beta)}\left(X_i, \hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i)\right) \tag{19}$$

It is worth noting that we use CATE estimates as they can only benefit the performance from a practical perspective. In general, however, policy learning approaches can still learn policies only with X_i as covariates. We repeat the process for any $\beta \in \mathcal{B}$ and obtain the set of policies $\Pi^{\text{MO-SPL}}$, which contains 501 policies.

Mixed Strategy Single Objective (`SO-Mix`): As a benchmark set of policies, we define a mixed strategy between single-objective policies. As such, for any probability $\alpha \in [0,1]$, we use the single-objective personalized policy for Ad Consumption with probability α and the single-objective personalized policy for Video Consumption with probability $1-\alpha$. For any α, we define policy $\pi_\alpha^{\text{SO-Mix}}$ as follows:

$$\pi_\alpha^{\text{SO-Mix}} = \mathbb{1}(Z_i \leq \alpha)\pi^{\text{SO-AC}}(X_i) + \mathbb{1}(Z_i > \alpha)\pi^{\text{SO-VC}}(X_i), \tag{20}$$

where Z_i is a random variable drawn from the uniform distribution: $Z_i \sim U(0,1)$. Intuitively, this pol-

icy covers any point on the line from $(\rho_A(\pi^{\text{SO-AC}}), \rho_V(\pi^{\text{SO-AC}}))$ to $(\rho_A(\pi^{\text{SO-VC}}), \rho_V(\pi^{\text{SO-VC}}))$. This serves as a great benchmark to see how much value adopting a multi-objective value can generate. We use $\alpha \in \{0, 1/500, 2/500, \ldots, 1\}$ to generate the set of mixed-strategy single objective policies $\Pi^{\text{SO-Mix}}$, which contains 501 policies.

Random Policy (`Random`): This set of policies is very similar to the Mixed Strategy Single Objective (`SO-Mix`) policies with a notable difference: instead of mixing between single-objective personalized policies, we mix between the uniform policies. For any $\alpha \in [0, 1]$, we use a mixed-strategy policy that uses the treatment condition (Skippable/Long) with probability α and the control condition (Non-Skippable/Short) with probability $1 - \alpha$. We denote the random policy with probability α by $\pi_\alpha^{\text{Random}}$ and formally define it as follows:

$$\pi_\alpha^{\text{Random}} = \mathbb{1}(Z_i \leq \alpha)1 + \mathbb{1}(Z_i > \alpha)0, \tag{21}$$

where 1 refers to the treatment assignment and 0 refers to the control assignment and $Z_i \sim U(0, 1)$ as before. Again, we use $\alpha \in \{0, 1/500, 2/500, \ldots, 1\}$ to generate the set of random policies Π^{Random} that contains 501 policies.

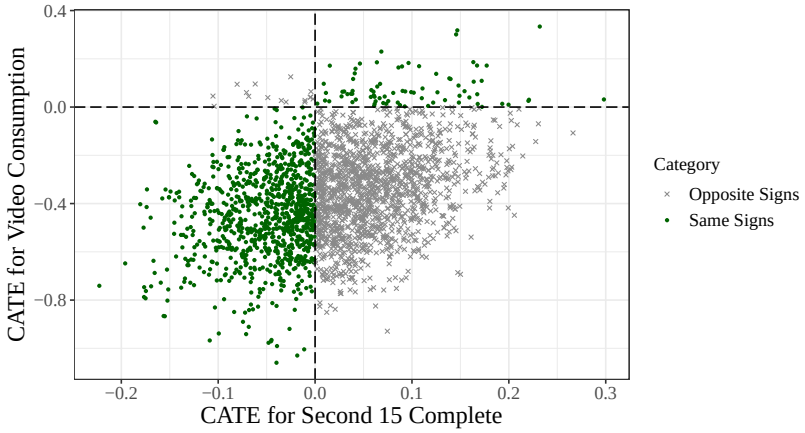
G Multi-Objective Personalization with Different Outcomes

Our main analysis focused on Ad Consumption as our ad-related metric. In the context of video ads, there can be substantial heterogeneity in what part of sponsored content consumption is more desired for both advertisers and platforms (Teixeira et al. 2014). An important feature of our framework is that we can apply it to any well-defined outcome. For video advertising platforms, another outcome is often of great interest: whether the user has reached the payment point. Platforms often set cutoff rules at 15 or 30 seconds to charge advertisers if the user reaches that point. One ad-related variable we can examine in our empirical setting is *Second 15 Complete*. In this section, we perform the multi-objective personalization framework when optimizing *Second 15 Complete* and *Video Consumption*. This approach reflects the joint utility of many video advertising platforms that charge advertisers when the user reaches a certain point within the ad. That is, if the user reaches a certain point in the ad, the advertiser has to pay even if the user later skips the ad. The cutoff rule varies across platforms, ranging from 15 to 30 seconds. Thus, a natural problem objective for platforms is to maximize the ad revenue by having more people reach the cutoff point while keeping Video Consumption high.

In Table 2, we presented the average treatment effect on Second 15 Complete as the outcome variable. We showed that users in the Skippable/Long ad condition are more likely to reach the 15^{th} second of the ad. However, the magnitude of the treatment effect is smaller compared to the effects on Ad Consumption. In particular, we do not expect a natural substitution pattern between Second 15 Complete and Video Consumption. Therefore, we expect that applying multi-objective personalization to this problem creates more substantial value.

We first estimate the CATE on Second 15 Complete using Causal Forests, with a 10-fold cross-validation. We plot the CATE on Video Consumption against CATE on Second 15 Complete for a random sample of observations in our and present the results in Figure A5 to see the extent to which the two outcomes are in

Figure A5. Scatter Plot of CATE Estimates for Video Consumption and Second 15 Complete



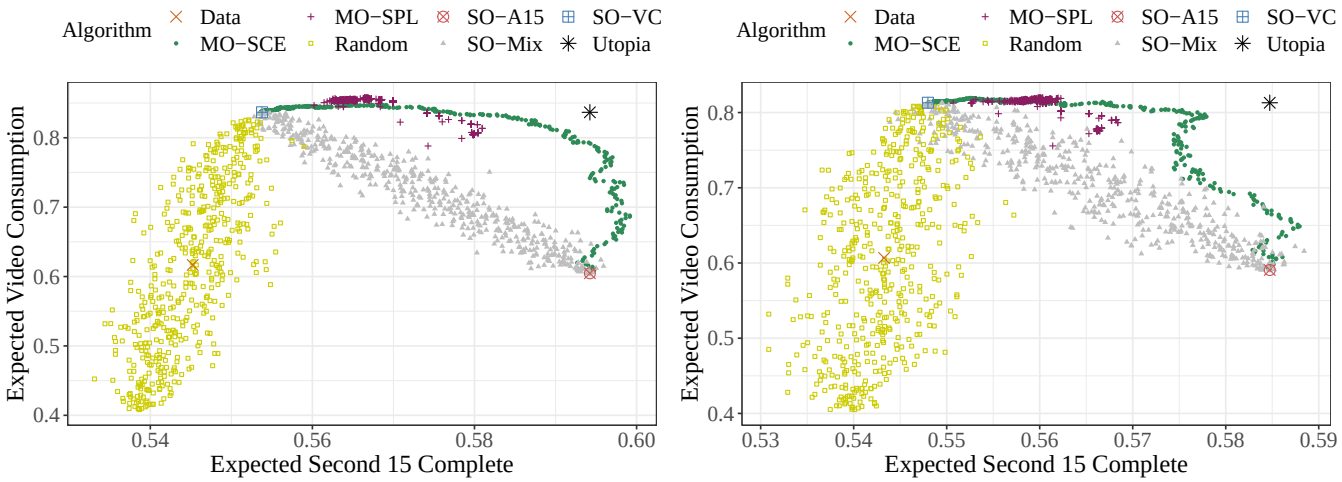
conflict. Unlike the case with Ad Consumption and Video Consumption, we note that the sign of the CATE estimates is the same for a large portion of observations. Over 41% of all observations have the same signs of CATE estimates, which means that the treatment assignment for these observations is clear: units with positive CATE estimates on both outcomes receive the Skippable/Long ad, whereas units with negative CATE estimates on both outcomes receive Non-Skippable/Short ad. Further, we find a positive correlation of 0.38 between CATE on Second 15 Complete and CATE on Video Consumption. This confirms our initial intuition that multi-objective personalization would be valuable in this setting.

We then apply both Scalarization with Causal Estimates (`MO-SCE`) and Scalarization with Policy Learning (`MO-SPL`) algorithms to generate a set of policies under each algorithm. We consider the same set of benchmark policies and references presented in §5.2: (1) Single-Objective for Second 15 Complete (`SO-A15`), (2) Single-Objective Video Consumption (`SO-VC`), (3) Mixed Strategy Single Objective (`SO-Mix`), (4) Random Policy (`Random`), (5) Data, and (6) Utopia.

We present the performance of all these policies using the IPS estimator when evaluated on the training and test data in Figure A6. A few insights immediately emerge from this figure. First, we note that the `MO-SCE` is more stable and covers a broader range of policies. Second, we find that the performance of `MO-SCE` is remarkably well on the training data as the Pareto frontier almost approaches the Utopia point. We then turn to our *Covered Area Proportion (CAP)* measure proposed in §3.3 to evaluate the performance of the identified sets of policies. We find that the set of `MO-SCE` policies achieve 75% and 58% in *CAP* measure on the training and test data, respectively. The performance is worse for the set of `MO-SPL` policies because their coverage is more sparse: the set of `MO-SPL` policies achieve 51% and 43% in *CAP* measure on the training and test data, respectively.

Third, we find policies on the identified Pareto frontier with a reasonable performance relative to both single-objective policies. In our case in the main text with Video Consumption and Ad Consumption as outcomes, the extensive substitution between the two outcomes prevented us from finding a single policy on the Pareto frontier that compares well with both single-objective policies. In the multi-objective personalization problem with Second 15 Complete and Video Consumption as outcomes, we find policy $\pi^{\text{MO-SCE}}_{0.760}$ from the set of `MO-SCE` policies that achieve great performance when compared to either single-objective

Figure A6. Policy Evaluation on Training and Test Data for Video Consumption and Second 15 Complete



(a) Training Data for Evaluation

(b) Test Data for Evaluation

policy. Compared to the Single-Objective Second 15 Complete (`SOA15`) policy, this multi-objective personalized policy improves Video Consumption by 36.4%, while reducing the Second 15 Complete rate by 1.1% on the training data. On the test data, the gain in Video Consumption is 35.1%, at the same loss of 1.2% of the Second 15 Complete rate. Interestingly, compared to the Single-Objective Video Consumption policy (`SO-VC`), the very same policy $\pi_{0.760}^{\text{MO-SCE}}$ increases the Second 15 Complete rate by 6.1% relative to the SOVC policy, while reducing Video Consumption by 1.4%. On the test data, the gain in the Second 15 Complete rate is 5.4%, and the drop in Video Consumption is 1.9%. Together, our results show that the multi-objective personalization framework can be applied to a variety of settings and generate gains beyond the single-objective personalization framework.

References

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

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.

Censor, Y. (1977). Pareto optimality in multiobjective problems. *Applied Mathematics and Optimization*, 4(1):41–59.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., and Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in medicine*, 32(19):3388–3414.

Pashkevich, M., Dorai-Raj, S., Kellar, M., and Zigmond, D. (2012). Empowering online advertisements by empowering viewers with the right to choose: the relative effectiveness of skippable video advertisements on youtube. *Journal of advertising research*, 52(4):451–457.

Súkeník, P. and Lampert, C. H. (2022). Generalization in multi-objective machine learning. *arXiv preprint arXiv:2208.13499*.

Teixeira, T., Picard, R., and El Kaliouby, R. (2014). Why, when, and how much to entertain consumers in advertisements? a web-based facial tracking field study. *Marketing Science*, 33(6):809–827.