



## Marketing Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Multiobjective Personalization of Marketing Interventions

Omid Rafieian, Anuj Kapoor, Amitt Sharma

To cite this article:

Omid Rafieian, Anuj Kapoor, Amitt Sharma (2024) Multiobjective Personalization of Marketing Interventions. Marketing Science

Published online in Articles in Advance 11 Sep 2024

. <https://doi.org/10.1287/mksc.2023.0122>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2024, INFORMS

Please scroll down for article—it is on subsequent pages





With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# Multiobjective Personalization of Marketing Interventions

Omid Rafieian,<sup>a,\*</sup> Anuj Kapoor,<sup>b</sup> Amitt Sharma<sup>c</sup>

<sup>a</sup>Marketing, Cornell Tech and SC Johnson College of Business, Cornell University, New York, New York 10044; <sup>b</sup>Department of Marketing, University of Missouri, Columbia, Missouri 65203; <sup>c</sup>vdo.ai, Vancouver, British Columbia V5Y 0E3, Canada

\*Corresponding author

Contact: [or83@cornell.edu](mailto:or83@cornell.edu),  <https://orcid.org/0000-0001-8633-2302> (OR); [anujkapoor@gmail.com](mailto:anujkapoor@gmail.com),  <https://orcid.org/0000-0002-3139-5098> (AK); [amitt@vdo.ai](mailto:amitt@vdo.ai) (AS)

Received: March 22, 2023

Revised: April 3, 2024; July 11, 2024

Accepted: July 16, 2024

Published Online in *Articles in Advance*:  
September 11, 2024

<https://doi.org/10.1287/mksc.2023.0122>

Copyright: © 2024 INFORMS

**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 multiobjective personalization framework that identifies personalized policies to balance multiple objectives at the individual level. We apply our framework to a canonical example of multiobjective conflict between sponsored and organic content consumption outcomes. Partnering with [vdo.ai](https://vdo.ai), we conduct a field experiment and randomly assign users to the skippable/long and nonskippable/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 multiobjective personalized policies can significantly improve both sponsored and organic consumption outcomes over single-objective policies. We show that compared with a single-objective policy optimized for organic consumption, there exists a multiobjective policy that increases sponsored consumption by 61% at the expense of only a 4% decrease in organic consumption. Similarly, compared with the single-objective policy optimized for sponsored consumption, there is a multiobjective policy that increases organic consumption by 53% while decreasing sponsored consumption by just 15%.

**History:** Olivier Toubia served as the senior editor.

**Supplemental Material:** The online appendix and data files are available at <https://doi.org/10.1287/mksc.2023.0122>.

**Keywords:** multiobjective personalization • causal inference • machine learning • field experiments • video advertising

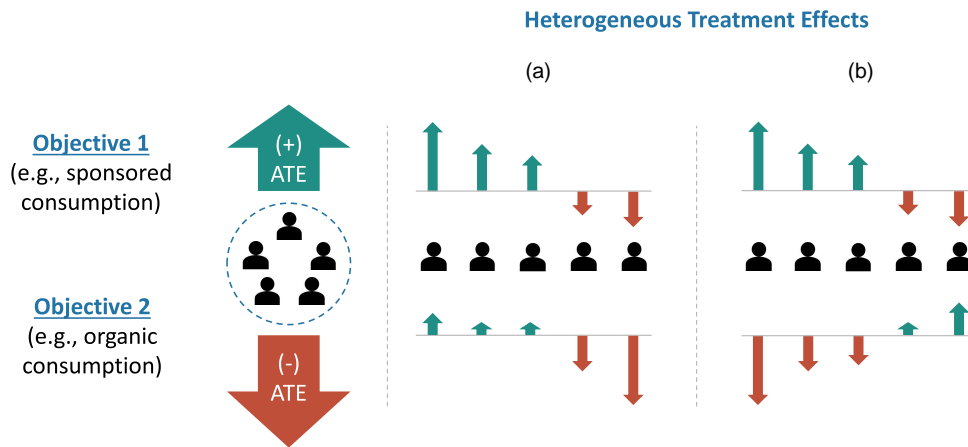
## 1. Introduction

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 higher organic content consumption to increase user engagement as well as 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 multiobjective 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 multiobjective 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 panel (b) of Figure 1.

In this paper, we study the problem of multiobjective 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 multiobjective personalization algorithms? What are the theoretical guarantees of these algorithms? How can we quantify the performance and the value created by multiobjective personalization algorithms? Next, we turn to a set of empirical questions. What are the gains from using a

**Figure 1.** (Color online) Overview of Multiobjective Personalization

Notes. Arrows show the treatment effects, and colors show the signs: green for positive and red for negative. Each individual is shown with an icon. In the aggregate case, they are grouped together. ATE, average treatment effect. (a) Optimistic scenario. (b) Pessimistic scenario.

multiobjective 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 or goal for our personalization problem. Unlike the single-objective personalization problem where the expected outcome values are totally ordered, the main challenge in multiobjective 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 multiobjective 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 multiobjective 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 multiobjective 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 multiobjective 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 multiobjective 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](https://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 nonskippable/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- and 15-second cuts of the same raw footage used by an actual advertiser. We only show preroll 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 multiobjective 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 nonskippable/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 more sponsored consumption on average compared with the nonskippable/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-point-lower organic consumption than the nonskippable/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, because 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-second increase in sponsored consumption results in approximately 13 seconds less 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 multiobjective personalization framework aims to address the substitution between sponsored and organic consumption. Intuitively, multiobjective 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 multiobjective 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 multiobjective personalization algorithms and this set of benchmarks quantifies the value that multiobjective 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 multiobjective personalization creates substantial value in this context. We document that compared with the single-objective personalized policy that only optimizes sponsored consumption, there is a multiobjective 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 with 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 multiobjective 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 multiobjective optimization to the literature on the intersection of machine learning and causal inference, and we propose a framework for multiobjective personalization. We prove that scalarization identifies the complete Pareto frontier in the outcome space. In particular, we propose two scalarization-based algorithms based on the heterogeneous treatment effect estimation and the policy learning that can identify the Pareto frontier of policies in terms of multiple objectives. From a managerial perspective, our multiobjective 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



multiobjective 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, Athey and Wager 2021, Nie 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 the software-as-a-service industry (Yoganarasimhan et al. 2023), and product versioning in music streaming platforms (Goli et al. 2024). The key insight in this series of work is that having a fine-grained set of pretreatment variables helps differentiate between users, thereby creating value by assigning users to the right policy. We extend this literature by proposing a multiobjective 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 multiobjective 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 nonmarketing problems where the manager needs to optimize more than one objective.

Second, our paper relates to the literature on multiobjective optimization (Marler and Arora 2004). This literature has proposed a series of algorithms to deal with multiobjective 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 multiobjective contextual bandits (Tekin and Turgay 2018; Turgay et al. 2018; Wang et al. 2022, 2024) and the literature on multiobjective reinforcement learning (Rojers et al. 2013, Van Moffaert and

Nowé 2014, Abdolmaleki et al. 2020). Our paper builds on the foundations laid out in the multiobjective 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 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, several papers have examined different aspects of ad avoidance by deriving the equilibrium properties in markets where users are averse to ads (Anderson and Coate 2005, Dukes et al. 2022), by quantifying the audience loss caused by ad avoidance (Wilbur 2008, Schweidel and Moe 2016, Rajaram et al. 2023), by predicting ad avoidance based on the past consumption of the product (Tuchman et al. 2018), by linking ad avoidance to sales (Bronnenberg et al. 2010, Deng and Mela 2018), and by proposing market design solutions to account for audience externalities (Wilbur et al. 2013).<sup>1</sup> 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).<sup>2</sup>

## 3. Multiobjective Personalization

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

### 3.1. Problem Definition

At a high level, multiobjective 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 empirical application. However, it is easy to extend our algorithms to multiple treatment settings. In a multiobjective personalization problem, we have  $K$  different outcomes denoted by  $\{(Y^{(1)}, Y^{(2)}, \dots, Y^{(K)})\}$ . 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  $\pi$  as follows.

**Definition 1.** A policy  $\pi: \mathcal{X} \rightarrow [0, 1]$  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  $\rho_j: \Pi \rightarrow \mathbb{R}$ , where  $\Pi$  is the space of all possible policies. This indicates that for a policy  $\pi$ , the performance in terms of outcome  $Y^{(j)}$  is characterized by  $\rho_j(\pi)$ . We can formally define this term as follows:

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

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  $\pi$ .

In a multiobjective personalization problem, we need to consider performance metrics for all outcomes of interest. As a result, comparing two policies is more challenging in a multiobjective case compared with 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  $\pi_1$  and  $\pi_2$  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

$\pi_1$  and  $\pi_2$ , such that  $\rho_s(\pi_1) > \rho_s(\pi_2)$  and  $\rho_o(\pi_1) > \rho_o(\pi_2)$ , we can conclude that policy  $\pi_2$  is dominated by  $\pi_1$  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 nondominated by any other policies. To this end, we define the main goal of multiobjective personalization as follows.

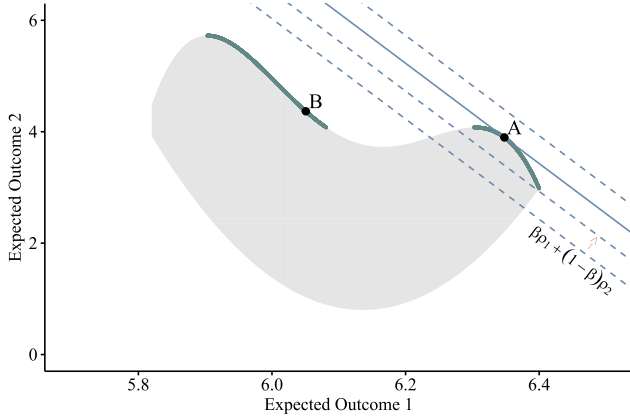
**Definition 3.** Suppose that there is a manager who wants to optimize multiple outcomes  $Y^{(1)}, Y^{(2)}, \dots, Y^{(K)}$ . Our goal is to find a set of policies  $\Pi_f$  that are Pareto optimal in terms of objectives  $\rho_1(\pi), \rho_2(\pi), \dots, \rho_K(\pi)$ . That is, for each  $\pi \in \Pi_f$ , there is no other policy  $\pi'$  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  $\Pi_f$ . In the next section, we discuss how we can design algorithms to identify these Pareto optimal policies.

### 3.2. Algorithms for Multiobjective Personalization

Fundamentally, the problem in Definition 3 can be viewed as a multiobjective optimization problem. The literature on multiobjective 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 is linked to multiple notions of reward or objective (Marler and Arora 2004). For example, a driver can set the continuous variables speed and total passenger weight to optimize the travel time and fuel cost. The multiobjective 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 multiobjective contextual bandits (Drugan and Nowe 2013; Tekin and Turgay 2018; Turgay et al. 2018; Wang et al. 2022, 2024) and multiobjective reinforcement learning (Roijsers 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 multiobjective 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 multiobjective optimization problem is linear scalarization, whereby we map multiple objectives into a single objective using linear weights that come from the probability simplex. That is, for any set of nonnegative

**Figure 2.** (Color online) Identification of Pareto Optimal Points Using Linear Scalarization

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 that for any given set of weights, the optimal solution  $\sum_{j=1}^K \beta_j \rho_j(\pi)$  is Pareto optimal.<sup>3</sup> 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 it fails 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 multiobjective personalization problem differs from the general class of multiobjective problems as we can formally prove that the outcome space  $\{(\rho_1(\pi), \dots, \rho_K(\pi))\}_{\pi \in \Pi}$  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  $\mathcal{B} = \{\beta \in \mathbb{R}_+^K, \sum_{j=1}^K \beta_j = 1\}$  denote the full set of linear weights. For any  $\beta \in \mathcal{B}$ , we define  $\pi_\beta^S = \arg \max_{\pi} \sum_{j=1}^K \beta_j \rho_j(\pi)$ , which is the solution to the single-objective problem with weights  $\beta$ . The full set of policies identified by linear scalarization, defined as  $\Pi^S = \{\pi_\beta^S\}_{\beta \in \mathcal{B}}$ , is the complete Pareto frontier of the outcome space (i.e.,  $\Pi^S = \Pi_f$ ).

**Proof.** See Online Appendix A.1 for the proof.  $\square$

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 that 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:

$$\pi_\beta^S = \arg \max_{\pi} \sum_{k=1}^K \beta_k \rho_k(\pi). \quad (2)$$

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  $\mathcal{D} = \{(X_i, W_i, Y_i^{(1)}, \dots, Y_i^{(K)})\}_{i=1}^N$ : (1) the stable unit treatment value assumption (SUTVA), which states that potential outcomes for one unit are not influenced by other units' treatment assignment and that there is only one version for each treatment; (2) the 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) the 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 because 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 because of the randomness in the data. Sùkenik 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 that our algorithms have for the single-objective case will conveniently transfer to the multiobjective 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 is denoted by  $\tau_k(\cdot)$  and defined as follows:

$$\tau_k(x) = \mathbb{E}[Y_i^{(j)}(1) - Y_i^{(j)}(0) | X_i = x]. \quad (3)$$

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  $\beta$ .

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

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

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

**Proof.** See Online Appendix A.2.  $\square$

The intuition behind Proposition 2 is the following; the only policy-variant element of  $\rho_j$  is  $\tau_j$ . 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  $\beta$  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 (Algorithm 1) precisely does that; it forms the union of all policies generated by different values of  $\beta \in \mathcal{B}$ .

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 Online Appendix B.1, we present a very simple illustrative example with three data points.

**Algorithm 1** (Scalarization with Causal Estimates)

**Input:**  $\{(\hat{\tau}_1(X_i), \hat{\tau}_2(X_i), \dots, \hat{\tau}_K(X_i))\}_{i=1}^N, \mathcal{B}$

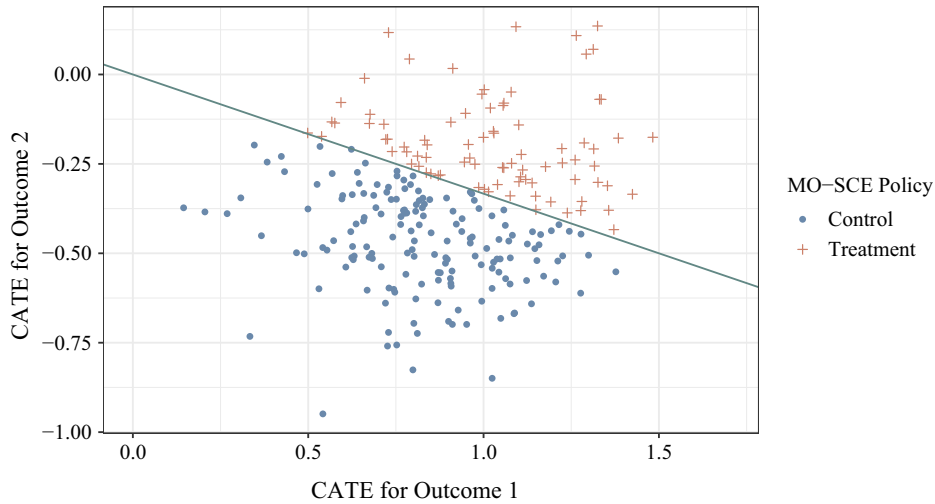
**Output:**  $\Pi^{\text{MO-SCE}}$

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

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 offer statistical guarantees for Algorithm 1. More specifically, the scalarization with causal estimates algorithm (Algorithm 1) performs well when the CATE estimates are more accurate.

**3.2.2. Scalarization with Policy Learning.** An alternative empirical approach to identify policies that

**Figure 3.** (Color online) An Illustration of the Assignment Policy in Scalarization with Causal Estimates



*Notes.* Each point shows the CATE on outcomes 1 and 2 for a data point. The figure illustrates the scalarization with the 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 signs), and points below that line are assigned to the control condition (blue circles).



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  $\pi$  from a certain policy class  $\Pi^*$  that optimizes the objective function. In this section, we want to define this approach for the multiobjective 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 estimator for outcome  $Y_i^{(j)}$  under policy  $\pi$  is defined as follows:

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

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  $\hat{\rho}_k^{\text{IPS}}(\cdot)$  is an unbiased estimator for  $\rho_j(\cdot)$ . As such, we can define the empirical equivalent of Equation (2) by plugging in  $\hat{\rho}_j^{\text{IPS}}(\cdot)$  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  $\mathcal{D} = \{(X_i, W_i, e(X_i), Y_i^{(1)}, \dots, Y_i^{(K)})\}_i$ . The variable  $\Gamma_i$  is defined at the individual

level as follows:

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

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

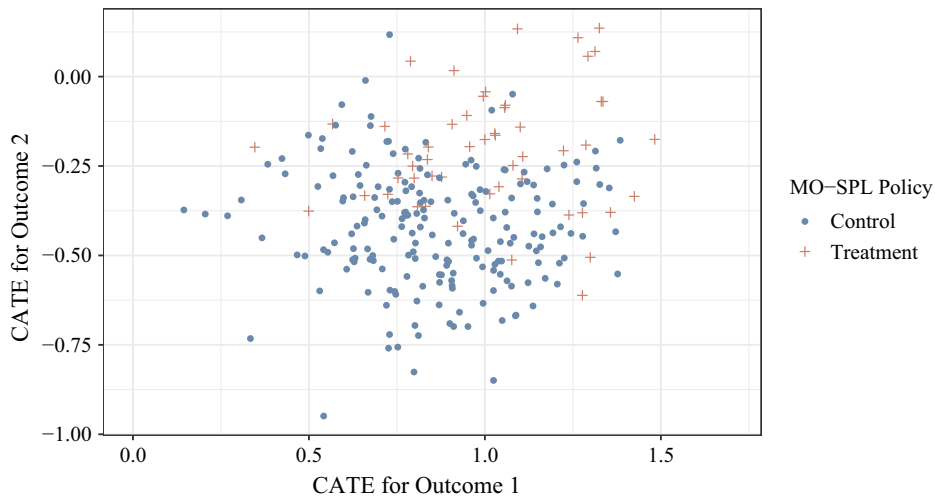
**Proof.** See Online Appendix A.3.  $\square$

This proposition indicates that for any  $\beta \in \mathcal{B}$ , we can directly learn policy  $\pi_\beta^S$  in Equation (2) using any off-the-shelf classification algorithm. The intuition behind this is the following; a more positive  $\Gamma_i$  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. It is worth emphasizing that the reason we estimate a policy function based on inputs  $X_i$  is that we do not know  $\Gamma_i$  prior to the treatment assignment.

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.<sup>4</sup>

We present the algorithm for multiobjective personalization that uses the policy learning approach in Algorithm 2. Like Algorithm 1, we obtain an estimate of policy  $\pi_\beta^S$  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

**Figure 4.** (Color online) An Illustration of the Assignment Policy in Scalarization with Policy Learning



*Notes.* Each point shows the CATE on outcomes 1 and 2 for a data point. The figure illustrates the scalarization with the policy learning policy for  $\beta = (0.25, 0.75)$ . The treatment (red plus signs) and control (blue circles) assignments are determined based on the weighted classification task.

present a very simple illustrative example with three data points in Online Appendix B.2.

#### Algorithm 2 (Scalarization with Policy Learning)

**Input:**  $\{X_i, W_i, e(X_i), Y_i^{(1)}, \dots, Y_i^{(K)}\}_i, \mathcal{B}$

**Output:**  $\Pi^{\text{MO-SPL}}$

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

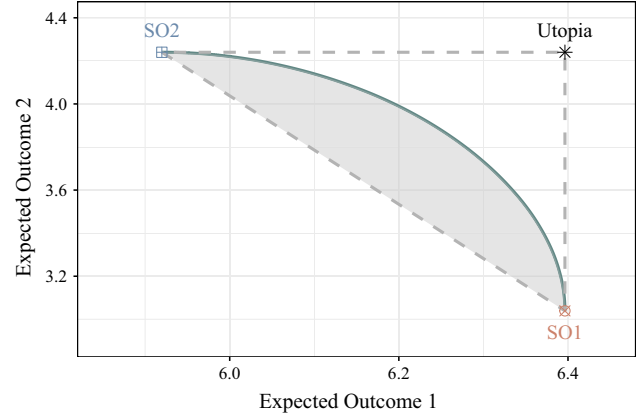
Like our scalarization with causal estimates algorithm (Algorithm 1), we require the standard assumptions of causal inference presented in Assumption 1 for identification. However, one key advantage of this approach is that even without theoretical guarantees on the  $1/\sqrt{N}$ -rate estimation of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

Suppose we have two algorithms  $A$  and  $B$  that identify policies  $\pi_A$  and  $\pi_B$  in a single-objective personalization problem. If we have an evaluator  $\hat{\rho}(\cdot)$ , we can compare the performance of  $\pi_A$  and  $\pi_B$  and see which one performs better. In a multiobjective personalization setting, however, algorithms  $A$  and  $B$  identify a set of policies  $\Pi_A$  and  $\Pi_B$ , 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 two-dimensional example in Figure 5 to illustrate our proposed measure. Each point in Figure 5 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

**Figure 5.** (Color online) A Visual Illustration of Multiobjective Policy Evaluation



*Notes.* 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 represents the optimal values in each dimension. The green curve is the Pareto frontier under a multiobjective algorithm. The covered area proportion measure is the proportion of the shaded area to the triangle between SO1, SO2, and utopia.

personalized policies thus serves as a reasonable benchmark for multiobjective personalization. On the other hand, the best-case scenario for multiobjective personalization is to achieve a *utopia* point as shown in Figure 5, which takes the optimal expected outcome in each dimension. The Pareto frontier will naturally be the 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 utopia, the greater the value created by multiobjective 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 multidimensional 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 it 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 multiobjective

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 that 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](https://vdo.ai) and conduct a field experiment to measure the 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 multiobjective 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](https://vdo.ai), which is based in India and the United States and provides video services to publishers worldwide. Since its launch, [vdo.ai](https://vdo.ai) has attracted many large- and medium-sized media publishers that use the company's technology to serve organic and sponsored video content on their web pages. As a form of monetization, [vdo.ai](https://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) preroll ads that are placed prior to the start of the video, (2) midroll ads that are placed in the middle of the video, and (3) postroll 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 preroll 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 that a user is visiting a web page. For example, the organic video content may be placed in a news article that a user wants to read. This is a common practice among media publishers, such as [cnn.com](https://cnn.com) or [espn.com](https://espn.com), that display organic and sponsored content within the news article. As the main facilitator of sponsored and organic video content, [vdo.ai](https://vdo.ai) wants

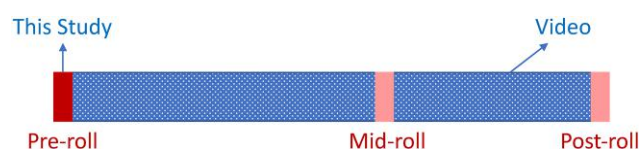
higher sponsored consumption as it is directly linked to ad revenues<sup>5</sup> as well as higher organic consumption as it relates to user engagement and attracts more content creators, which help the platform generate more advertising opportunities in the short run and the long run.

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.<sup>6</sup> 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](https://vdo.ai), where the preroll impressions are assigned to three experimental conditions: (1) the no-ad condition, where the user does not need to watch an ad to start consuming the organic video content; (2) the non-skippable/short ad condition, where the ad shown is a 15-second-long ad of boAt's Watch Xtend product that is not skippable; and (3) the skippable/long ad condition, where the ad shown is a 60-second-long ad of the same boAt product that is skippable after 5 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 that 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 Online Appendix C, we present more details about the advertised product and brand, and we 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 the 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.

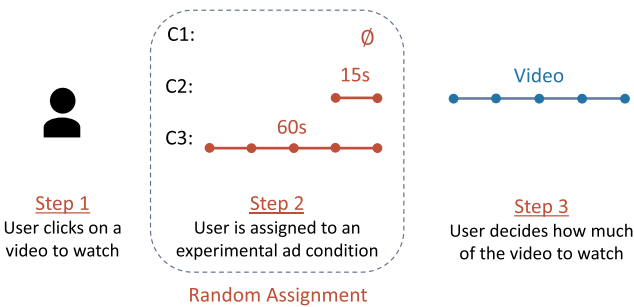
Figure 7 summarizes a schema of our experiment and the treatment conditions. The points in Figure 7 show the pixels placed to find whether the user has reached a certain point in the ad and video. This means that we can record whether the user has reached the midpoint of the skippable/long ad (i.e.,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

**Figure 6.** (Color online) Different Types of Video Ads





**Figure 7.** (Color online) A Visual Schema of the Experiment Design



*Notes.* 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. s, seconds. C1-C3 refers to Condition1-Condition3.

video content 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 *pretreatment variables*, *treatment assignment*, and a rich set of *posttreatment variables* or *outcomes* as the main focus of our empirical application is to study policy-making when there are many outcomes. For pretreatment variables, we observe the user’s *Internet Protocol Address*, *Time of Day*, *Date*, *City*, and *Country*, and we observe the *Operating System (OS)* of the device that the user is using. For posttreatment variables or outcomes, we collect a rich set of posttreatment variables or outcomes both on the ad performance and on the 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,<sup>7</sup> and for the video, these pixels are placed at every quarter (25%) of the video.<sup>8</sup> In addition to the pixels shown in Figure 7,

we also 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 is relatively low.<sup>9</sup> 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 the 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, because 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 reduce our sample size to 53,176 sessions generated by 51,423 unique users. Because 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 pretreatment 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 Table 1, the hours with the highest traffic are 6–8 a.m. Mountain Standard Time (MST), which would be 5:30–7:30 p.m. 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

**Table 1.** Summary Statistics of the Pretreatment Variables

Variable	No. of subcategories	Top three subcategories and their shares		
		First	Second	Third
Hour of day	24	7 a.m. MST (7.76%)	6 a.m. MST (6.82%)	8 a.m. 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%)



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 pretreatment variable show that Android OS is the most common OS in our data, with around 80% of the total traffic. In Online Appendix D, we perform extensive randomization checks on the distribution of pretreatment 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 we let  $X$  and  $W$  denote the pretreatment covariates and the treatment variable, respectively. Because 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 nonskippable/short ad formats, respectively.<sup>10</sup> For inference, we use the common assumptions in the causal inference literature: (1) overlap, (2) unconfoundedness, and (3) the stable unit treatment value assumption. The first two are satisfied by design because 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 multiobjective personalization framework, and we present a detailed description of outcomes and an interpretation of the results in Online Appendices E.1 and E.2, respectively.

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 upper part of Table 2 presents the average treatment

effects for ad-related outcomes. We find that the average consumption of the skippable/long ad is significantly higher than that of the nonskippable/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 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 nonskippable/short version, despite the skippability of the ad in this condition. When we focus on the ad complete outcome, we find that the nonskippable/short ad has a substantially higher completion rate, which is expected because of the shorter length and nonskippability 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 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 nonskippable/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 nonskippable/short condition than in 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

**Table 2.** Average Treatment Effects Across Outcomes

Outcome	Mean of treatment A (nonskippable/short)	Mean of treatment B (skippable/long)	Mean difference $B - A$ estimate	$p$ -value
Ad consumption ( $\times 15$ seconds)	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.

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

	Outcome: Video consumption (quarters)	
	(1) Ordinary least squares	(2) Instrumental variable (2SLS)
Ad consumption (×15 seconds)	0.3538*** (0.0042)	−0.4209*** (0.0150)
Instruments	None	Treatment
Weak instruments test		7,567***
No. of observations	51,423	51,423

Notes. Standard errors are reported in parentheses. The weak instrument hypothesis is rejected as the  $F$ -statistic in the first-stage regression is 7,567 ( $p < 0.001$ ). 2SLS, two-stage least-squares.

\*\*\* $p < 0.001$ .

approach is to regress video consumption on ad consumption to see how the two outcomes are linked. 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 in column (1) in Table 3 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) in 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. Although we do not have the information about the exact length of videos, we know that the average is around two 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  $120 \times 0.1052 = 12.62$  seconds, 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 multiobjective personalization algorithms.

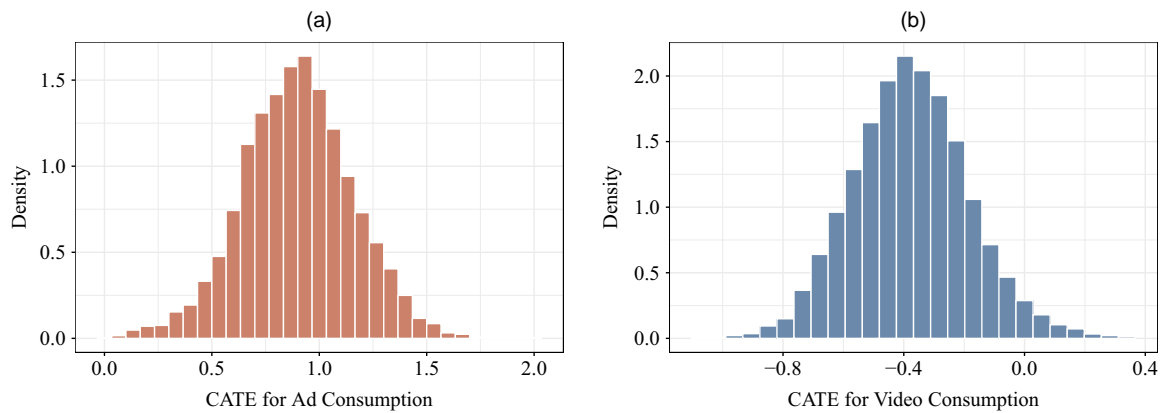
#### 4.4. Heterogeneous Treatment Effect Analysis

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 defined in Section 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  $\mathcal{I}_a \cap \mathcal{I}_v = \emptyset$ , 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) and Athey et al. (2019) for detailed algorithm presentation. For the set of covariates, we use all of the pretreatment variables presented in Table 1 as well as the exact time stamp to capture more fine-grained time-dependent heterogeneity and the latitude and longitude of cities to go beyond the city categories and capture the spatial heterogeneity patterns (if any). We use 10-fold crossvalidation to tune the hyperparameters of the causal forest.

We present the histogram of our CATE estimate for both ad consumption and video consumption outcomes. Figure 8(a) shows how CATE for the ad consumption outcome varies across individuals. As shown in Figure 8(a), 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.

**Figure 8.** (Color online) The Distribution of CATE Estimates for Ad and Video Consumption

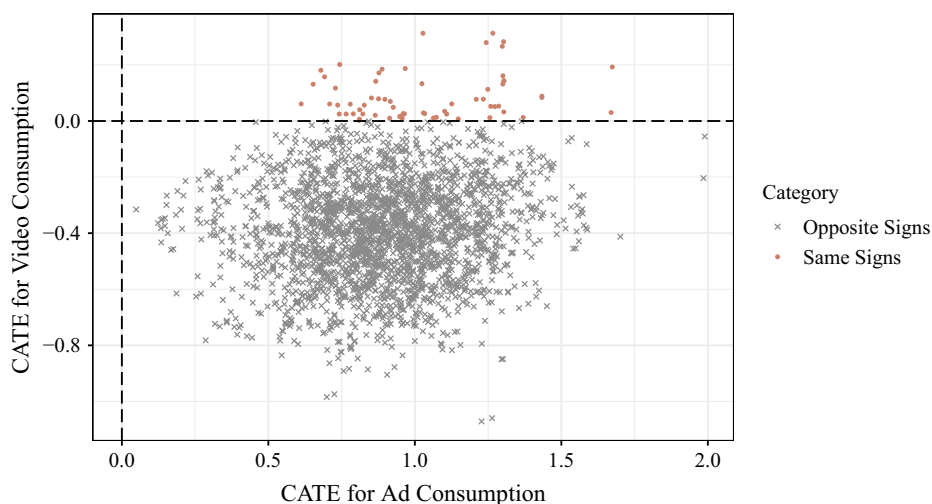
Notes. (a) Outcome: ad consumption ( $\times 15$  seconds). (b) Outcome: video consumption (in quarters).

We then move on to CATE estimates for video consumption as our video-related outcome and visualize the distribution of CATE estimates in Figure 8(b). As shown in Figure 8(b), 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 policy where all users are assigned to a nonskippable/short ad. For both outcomes, we present a more interpretable analysis of treatment effect heterogeneity in Online 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

scatterplot for a random sample of our observations in Figure 9. The first pattern that emerges from Figure 9 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 circles 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? Because 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

**Figure 9.** (Color online) Scatterplot of CATE Estimates on Video Consumption and Ad Consumption

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 multiobjective 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 Multiobjective Personalization

### 5.1. Counterfactual Policy Evaluation

Our main algorithms in Section 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  $\pi$  becomes one of counterfactual policy evaluation. Because our CATE estimates are structural parameters, we can relatively compare the performance of a policy  $\pi$  with any given baseline policy. Let  $\bar{Y}^{(j)}$  denote the average observed for outcome  $j$  in the data. We can write

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

where the elements of this sum are only nonzero 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 estimator defined in Definition 5 as follows:

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

IPS provides an unbiased estimator of the expected outcome under any policy  $\pi$ . 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  $\pi$  been implemented.<sup>11</sup> 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. Because the IPS estimator is defined on the data  $\mathcal{D}$ , we can easily evaluate both the in-sample performance and the 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 policymaking, 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 follows, and we refer the reader to Online Appendix F for greater details.

- *Scalarization with causal estimates.* We use  $\mathcal{B} = \{(\beta, 1 - \beta) | \beta = i/500, 0 \leq i \leq 500\}$  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.* We use the same  $\mathcal{B}$ , but we 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.<sup>12</sup> 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  $\alpha$  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  $\alpha$  and the control condition (nonskippable/short) with probability  $1 - \alpha$ . For consistency, we use  $\alpha \in \{i/500\}_{i=0}^{500}$  to generate the set  $\Pi^{\text{Random}}$ .

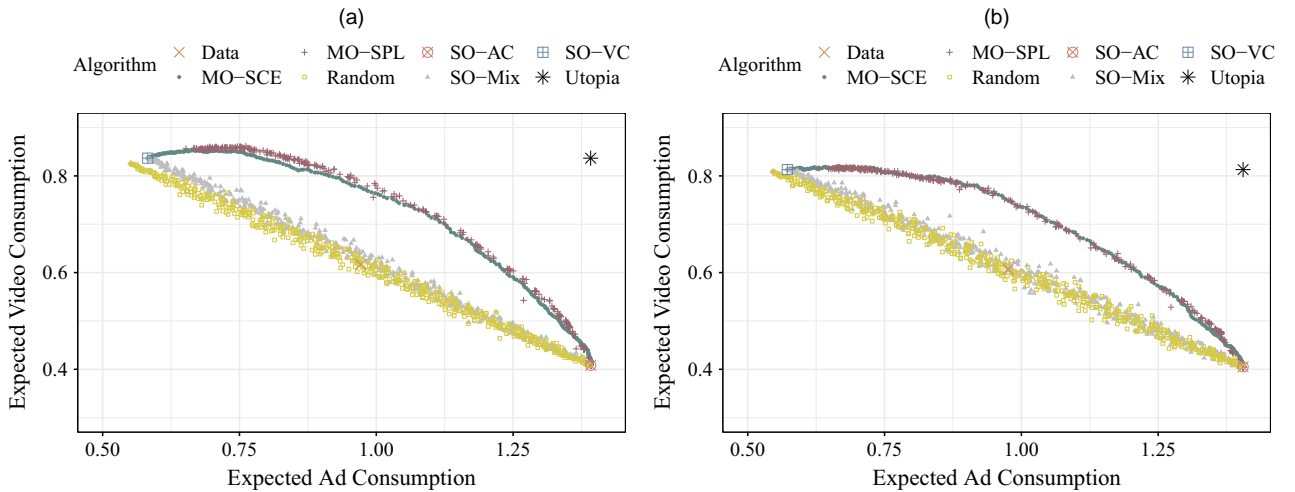
Overall, we have four sets of policies ( $\Pi^{\text{MO-SCE}}$ ,  $\Pi^{\text{MO-SPL}}$ ,  $\Pi^{\text{SO-Mix}}$ , and  $\Pi^{\text{Random}}$ ) and two policies ( $\pi^{\text{SO-AC}}(\cdot)$  and  $\pi^{\text{SO-VC}}(\cdot)$ ). The single-objective uniform policies and the fully random policies are also 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.

### 5.3. Policy Comparison Results

For any policy  $\pi$ , 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  $(\hat{\rho}_A^{\text{IPS}}(\pi; \mathcal{D}_{\text{Train}}), \hat{\rho}_V^{\text{IPS}}(\pi; \mathcal{D}_{\text{Train}}))$  and  $(\hat{\rho}_A^{\text{IPS}}(\pi; \mathcal{D}_{\text{Test}}), \hat{\rho}_V^{\text{IPS}}(\pi; \mathcal{D}_{\text{Test}}))$  in training data 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 in Figure 10 constitute the Pareto frontier under the scalarization with causal estimates algorithm, whereas purple plus signs in Figure 10 constitute the Pareto frontier under the scalarization with policy learning algorithm.<sup>13</sup> Because 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 nonuniform.

When comparing these Pareto frontiers with the single-objective policies, we see substantial gains as most mixed strategy single-objective 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* measure defined in Section 3.3, and we calculate the proportion of the area covered between the SO-Mix and utopia. On the training set, we find 43% and 48%

**Figure 10.** (Color online) Policy Evaluation on Training and Test Data



*Notes.* 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 for ease of exposition. (a) Training data for evaluation. (b) Test data for evaluation.

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 multiobjective 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 multiobjective personalization focuses on identifying 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 multiobjective personalization is in providing a complete picture for managers so that they can 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:

**5.3.1. High Video Consumption, Medium Ad Consumption.** From Figure 10(a), 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 with the SO-VC policy while increasing ad consumption by 60.0% on the training data. When comparing the performance of this policy ( $\pi_{0.246}^{\text{MO-SCE}}$ ) 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).

**5.3.2. High Ad Consumption, Medium Video Consumption.** On the right of Figure 10(a), 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 with the single-objective ad consumption policy while only losing 13.3% in the expected ad consumption. On the test data, the policy  $\pi_{0.420}^{\text{MO-SPL}}$  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 midroll ad. We use the IPS estimator for this outcome to evaluate the proportion of users who reach

that point and generate a midroll impression under both  $\pi_{0.420}^{\text{MO-SPL}}$  and SO-AC. Interestingly, we find that the proportion of users who reach the first quarter of the video is 22.7% under  $\pi_{0.420}^{\text{MO-SPL}}$ . In contrast, this proportion is 14.7% under SO-AC, suggesting a 55.5% increase in the number of potential midroll ad impressions that would be generated under  $\pi_{0.420}^{\text{MO-SPL}}$  relative to SO-AC.

**5.3.3. High Video Consumption, 15-Second Ad Consumption.** A useful feature of multiobjective personalization is that we can fix a value for one objective and examine the performance in terms of the other objective. Because ad consumption cannot technically be more than 15 seconds in the nonskippable/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  $\pi_{0.274}^{\text{MO-SCE}}$  with the two single-objective policies. Compared with the single-objective video consumption 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  $\pi_{0.274}^{\text{MO-SCE}}$  improves video consumption by 81.9% compared with the single-objective ad consumption policy while losing 29.2% in ad consumption.

In summary, we find that multiobjective personalization results in substantial gains in one objective without sacrificing too much in the other objective. Intuitively, multiobjective 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 multiobjective 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 Online Appendix G: (1) second 15 complete and (2) video consumption. Because many platforms charge advertisers once their ad is watched for 15 seconds (e.g., Facebook), using these two objectives for multiobjective personalization is reasonable for profit-maximizing platforms. We present the results of this practice in Online Appendix G and document substantial gains from a multiobjective 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 multiobjective 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 of the dominated policies. Further, platforms can identify the Pareto optimal policies on a batch of training data and use them 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 we showed that the platform could create value using our multiobjective 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 multiobjective personalization for *second 15 complete* and *video consumption* as our main outcomes of interest (see Online 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 should also be based on the externality that 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 second 30 of the ad. Part of the reason for having these rules in place is to account for the externalities that 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 that 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 multiobjective personalization framework that can reliably identify the Pareto frontier of personalized policies in terms of multiple outcomes. In particular, we propose the two algorithms *scalarization with causal estimates* and *scalarization with policy learning* that combine the insights from the causal machine learning literature



with that of multiobjective 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 versus organic content consumption. In collaboration with [vdo.ai](#), we conducted randomized experiments where users were randomly assigned to either the skippable/long or nonskippable/short versions of the same ad. We document a high degree of substitution between the two key outcomes *ad consumption* and *video consumption*, even at the individual level. We then apply our multiobjective personalization algorithms and find that the resulting policies improve the outcome in one dimension compared with single-objective personalized policies without sacrificing the outcome in the other dimension. In particular, we show that compared with 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 with the single-objective personalized policy that only optimizes ad consumption, there is a multiobjective 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 multiobjective optimization literature with causal machine learning and present a framework for multiobjective personalization. Our proposed algorithms take approaches for conditional average treatment effect estimation and policy learning, and they 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 multiobjective 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.

## Acknowledgments

The authors thank Vaibhav Mishra, Hemank Bajaj, Kushagra Gupta, and Krishnendu Pain for excellent research assistance on this project. The authors thank Sachin Gupta, Jura Liaukonyte, Hema Yoganarasimhan, and Dennis Zhang for detailed comments that have improved the paper. The authors also thank the participants of the Brown Bag at Cornell University; Centre for Data Science and Artificial Intelligence at Indian Institute of Management Ahmedabad; the Chicago Booth India Quantitative Marketing Conference 2023; and marketing seminars at Washington University in St. Louis, Duke, Koc University, the University of Bocconi, the University of Chile, and Wharton for their feedback. The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or nonfinancial interest in the subject matter or materials discussed in this manuscript.

## Endnotes

<sup>1</sup> See Wilbur (2016) for a great summary of the ad avoidance literature.

<sup>2</sup> We use sponsored consumption interchangeably with ad consumption throughout the paper, and we use organic consumption interchangeably with video consumption throughout the paper.



<sup>3</sup> We can prove it by contradiction; if there is a policy  $\pi'$  that dominates  $\pi$  in all objectives, then it is a contradiction that  $\pi$  is the optimal solution to the scalarized objective.

<sup>4</sup> 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 maps them to policies, which induce 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.

<sup>5</sup> 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.

<sup>6</sup> This inventory is unsold only for research and development purposes, and it 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.

<sup>7</sup> It is worth noting that the ad consumption in the nonskippable/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.

<sup>8</sup> 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 because we randomize the treatment.

<sup>9</sup> 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).

<sup>10</sup> We only use these two treatment groups in the main text, but we present some results with the no-ad condition in Online Appendix E.

<sup>11</sup> See Rafieian and Yoganarasimhan (2023) for a detailed explanation of the intuition behind this estimator.

<sup>12</sup> Normally, one could drop the CATE estimates from inputs. However, the model with estimated CATEs ensures a better performance, so we include them as inputs.

<sup>13</sup> 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.

## References

- Abdolmaleki A, Huang S, Hasenclever L, Neunert M, Song F, Zambelli M, Martins M, Heess N, Hadsell R, Riedmiller M (2020) A distributional view on multi-objective policy optimization. *Internat. Conf. Machine Learn.* (PMLR, New York), 11–22.
- Anderson SP, Coate S (2005) Market provision of broadcasting: A welfare analysis. *Rev. Econom. Stud.* 72(4):947–972.
- Ascarza E (2018) Retention futility: Targeting high-risk customers might be ineffective. *J. Marketing Res.* 55(1):80–98.
- Athey S, Wager S (2021) Policy learning with observational data. *Econometrica* 89(1):133–161.
- Athey S, Tibshirani J, Wager S (2019) Generalized random forests. *Ann. Statist.* 47(2):1148–1178.
- Belloni A, Lovett MJ, Boulding W, Staelin R (2012) Optimal admission and scholarship decisions: Choosing customized marketing offers to attract a desirable mix of customers. *Marketing Sci.* 31(4):621–636.
- Bronnenberg BJ, Dubé J-P, Mela CF (2010) Do digital video recorders influence sales? *J. Marketing Res.* 47(6):998–1010.
- Danaher PJ (1995) What happens to television ratings during commercial breaks? *J. Advertising Res.* 35(1):37–47.
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evolutionary Comput.* 6(2):182–197.
- Deng Y, Mela CF (2018) TV viewing and advertising targeting. *J. Marketing Res.* 55(1):99–118.
- Drugan MM, Nowe A (2013) Designing multi-objective multi-armed bandits algorithms: A study. 2013 *Internat. Joint Conf. Neural Networks (IJCNN)* (IEEE, Piscataway, NJ), 1–8.
- Dukes A, Liu Q, Shuai J (2022) Skippable ads: Interactive advertising on digital media platforms. *Marketing Sci.* 41(3):528–547.
- Ferguson S (2023) 8 Ways to increase brand awareness with video in 2023. Accessed February 10, 2023, <https://www.wyzowl.com/brand-awareness/>.
- Goli A, Reiley DG, Zhang H (2024) Personalized versioning: Product strategies constructed from experiments on pandora. *Management Sci.*, ePub ahead of print August 6, <https://doi.org/10.1287/mksc.2022.0357>.
- Horvitz DG, Thompson DJ (1952) A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* 47(260):663–685.
- Lambrecht A, Tucker C (2019) Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of stem career ads. *Management Sci.* 65(7):2966–2981.
- Marler RT, Arora JS (2004) Survey of multi-objective optimization methods for engineering. *Structural Multidisciplinary Optim.* 26:369–395.
- McGranaghan M, Liaukonyte J, Wilbur KC (2022) How viewer tuning, presence, and attention respond to ad content and predict brand search lift. *Marketing Sci.* 41(5):873–895.
- Miettinen K, Mäkelä MM (2002) On scalarizing functions in multi-objective optimization. *OR Spectrum* 24:193–213.
- Munro E (2024) Treatment allocation with strategic agents. *Management Sci.*, ePub ahead of print March 27, <https://doi.org/10.1287/mnsc.2022.01629>.
- Neyman J (1923) Sur les applications de la théorie des probabilités aux expériences agricoles: Essai des principes. *Roczniki Nauk Rolniczych* 10:1–51.
- Nie X, 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. Working paper, Cornell University, New York.
- Rafieian O (2023) Optimizing user engagement through adaptive ad sequencing. *Marketing Sci.* 42(5):910–933.
- Rafieian O, Yoganarasimhan H (2021) Targeting and privacy in mobile advertising. *Marketing Sci.* 40(2):193–218.
- Rafieian O, Yoganarasimhan H (2023) AI and personalization. Sudhir K, Toubia O, eds. *Artificial Intelligence in Marketing*, Review of Marketing Research, vol. 20 (Emerald Publishing Limited, Leeds, UK), 77–102.
- Rajaram P, Manchanda P, Schwartz EM (2023) Finding the sweet spot: Ad delivery on streaming media. Preprint, submitted April 22, <http://dx.doi.org/10.2139/ssrn.3496039>.
- Rojers DM, Vamplew P, Whiteson S, Dazeley R (2013) A survey of multi-objective sequential decision-making. *J. Artificial Intelligence Res.* 48:67–113.
- Schweidel DA, Moe WW (2016) Binge watching and advertising. *J. Marketing* 80(5):1–19.
- Shalit U, Johansson FD, Sontag D (2017) Estimating individual treatment effect: Generalization bounds and algorithms. *Internat. Conf. Machine Learn.* (PMLR, New York), 3076–3085.

- Siddarth S, Chattopadhyay A (1998) To zap or not to zap: A study of the determinants of channel switching during commercials. *Marketing Sci.* 17(2):124–138.
- Simester D, Timoshenko A, Zoumpoulis SI (2020a) Efficiently evaluating targeting policies: Improving on champion vs. challenger experiments. *Management Sci.* 66(8):3412–3424.
- Simester D, Timoshenko A, Zoumpoulis SI (2020b) Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Sci.* 66(6):2495–2522.
- Súkeník P, Lampert CH (2022) Generalization in multi-objective machine learning. Preprint, submitted August 29, <https://arxiv.org/abs/2208.13499>.
- Sun M, Zhu F (2013) Ad revenue and content commercialization: Evidence from blogs. *Management Sci.* 59(10):2314–2331.
- Swaminathan A, Joachims T (2015) Counterfactual risk minimization: Learning from logged bandit feedback. *Internat. Conf. Machine Learn.* (PMLR, New York), 814–823.
- Tekin C, Turgay E (2018) Multi-objective contextual multi-armed bandit with a dominant objective. *IEEE Trans. Signal Processing* 66(14):3799–3813.
- Tuchman AE, Nair HS, Gardete PM (2018) Television ad-skipping, consumption complementarities and the consumer demand for advertising. *Quant. Marketing Econom.* 16(2):111–174.
- Turgay E, Oner D, Tekin C (2018) Multi-objective contextual bandit problem with similarity information. *Internat. Conf. Artificial Intelligence Statist.* (PMLR, New York), 1673–1681.
- Van Moffaert K, Nowé A (2014) Multi-objective reinforcement learning using sets of pareto dominating policies. *J. Machine Learn. Res.* 15(1):3483–3512.
- Wager S, Athey S (2018) Estimation and inference of heterogeneous treatment effects using random forests. *J. Amer. Statist. Assoc.* 113(523):1228–1242.
- Wang Y, Tao L, Zhang XX (2024) Recommending for a multi-sided marketplace: A multi-objective hierarchical approach. *Management Sci.*, ePub ahead of print August 14, <https://doi.org/10.1287/mksc.2022.0238>.
- Wang N, Wang H, Karimzadehgan M, Kveton B, Boutilier C (2022) Imo <sup>^</sup>3: Interactive multi-objective off-policy optimization. Raedt LD, ed. *Proc. 31st Internat. Joint Conf. Artificial Intelligence* (IJCAI.org), 3523–3529.
- Wilbur KC (2008) A two-sided, empirical model of television advertising and viewing markets. *Marketing Sci.* 27(3):356–378.
- Wilbur KC (2016) Advertising content and television advertising avoidance. *J. Media Econom.* 29(2):51–72.
- Wilbur KC, Xu L, Kempe D (2013) Correcting audience externalities in television advertising. *Marketing Sci.* 32(6):892–912.
- Yoganarasimhan H, Barzegary E, Pani A (2023) Design and evaluation of optimal free trials. *Management Sci.* 69(6): 3220–3240.
- Zufryden FS, Pedrick JH, Sankaralingam A (1993) Zapping and its impact on brand purchase behavior. *J. Advertising Res.* 33(1): 58–67.