Multi-Objective Personalization of the Length and Skippability of Video Advertisements

Omid Rafieian*
Cornell University

Anuj Kapoor*
IIM Ahmedabad

Amitt Sharma
VDO.AI

March 22, 2023

Abstract

In this paper, we study two features of digital video ads on content-streaming platforms: length and skippability. Working with `vdo.ai`, we conduct a field experiment and randomly assign users to the Skippable/Long and Non-Skippable/Short versions of the same ad. We find that compared to the Non-Skippable/Short ad, the Skippable/Long ad version in our study increases ad consumption but decreases video consumption. This substitution pattern between ad and video consumption leads to a challenge for platforms seeking to maximize both outcomes. To address this challenge, we develop algorithms for multi-objective personalization that use individual-level substitution patterns to optimize ad and video consumption. The results show that multi-objective personalized policies can significantly improve both ad and video consumption outcomes over single-objective policies. In particular, we show that compared to a single-objective policy optimized for video consumption, there exists a multi-objective policy on the Pareto frontier that increases ad consumption by 61% at the expense of only a 4% decrease in video consumption. Similarly, compared to the single-objective policy optimized for ad consumption, there is a multi-objective policy that increases video consumption by 47% while decreasing ad consumption by just 13%. We conclude by discussing the practical implications for platform decision-making in real time.

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

*We would like to thank Vaibhav Mishra, Hemank Bajaj, and Kushagra Gupta for excellent research assitance on this project. We thank Sachin Gupta, Jura Liaukonyte, and Hema Yoganarasimhan for detailed comments that have improved the paper. Please address all correspondence to: or83@cornell.edu and anujk@iima.ac.in.

1 Introduction

Video advertising on content streaming platforms has become one of the dominant channels for advertisers. According to a recent report by Internet Advertising Bureau (IAB), digital video advertising is estimated to surpass $50 Billion in advertising spend in 2023 (IAB 2022). The popularity of digital video ads among marketers stems from the combination of great features of traditional TV advertising, such as audio-visual engagement, with those of digital ads, such as tracking, personalization, and better ad measurements. These advantages make digital video advertising an effective medium for advertisers and a sustainable business model for content-streaming platforms.

Digital video ads typically vary in two important features: length and skippability. Longer ads are generally more informative. However, these ads can result in ad avoidance, thereby negatively affecting the ad and the streaming platform. As a solution to the ad avoidance problem, platforms offer a skip option whereby users can skip the ad after a short period of time. Skippability can act as a screening process that filters users uninterested in the ad content. However, advertisers sometimes want to reach uninterested users, and skippability can hurt this goal. As such, many advertisers choose to make their ads non-skippable but short to minimize ad avoidance. Together, two common ad formats have emerged from this interplay between the length and skippability features: Non-Skippable/Short and Skippable/Long ad formats. Advertisers often create both versions of the same product with the same video material such that both versions have many scenes in common.

For platforms, it is not clear which one of these two ad formats is optimal. Part of the confusion comes from the presence of multiple objectives for the platforms. In particular, platforms have some target ad-related and video-related metrics, some of which are guided by their monetization strategy. For example, they want a higher engagement with both the sponsored ad content and the organic video content. However, these objectives are sometimes in direct conflict with each other. On the one hand, higher ad engagement can substitute for the time users spend on videos. On the other hand, one could argue that a lower ad avoidance and a higher engagement with the ad content can positively affect the engagement of the user with the video content. The platform's key challenge is finding a policy that achieves desirable outcomes in terms of both ad and video consumption.

In this paper, we study the substitution between ad and video consumption in online streaming platforms. We view the problem through the lens of a platform and address the following questions:

1. How does a Skippable/Long ad format perform in terms of ad and video consumption compared to the Non-Skippable/Short ad format? To what extent is there a substitution pattern between ad and video consumption in content-streaming platforms?

2. Is there any heterogeneity in the effect of the Skippable/Long ad format on both ad and video consumption outcomes compared to the Non-Skippable/Short ad format?

3. How can a platform develop a set of personalized policies that perform well in terms of multiple objectives? What are the gains of multi-objective personalization?

To answer these questions, we face three sets of challenges. First, to compare the performance of Skippable/Long ads with that of Non-Skippable/Short ads, we need to have exogenous variation in users' assignment to each condition. Second, to quantify the heterogeneity in the effects of these ad formats, we need a method to reliably estimate conditional average treatment effects (CATE). Third, to perform multi-objective personalization, we need a framework that finds a set of personalized policies that perform well on both dimensions (i.e., ad consumption and video consumption). In particular, our goal is to find the Pareto frontier of the policies in terms of both ad and video consumption objectives, i.e., the set of policies that are not dominated by any other policy in both objectives.

To address our first challenge, we partner with the video advertising platform `vdo.ai` and run a field experiment where we randomly assign users to different ad conditions. In particular, we assign users to three experimental conditions: (1) a Skippable/Long ad of a single product that is 60 seconds long, (2) a Non-Skippable/Short ad of the same product that is 15 seconds long, and (3) a No-Ad condition where the user watches the video content without having to watch an ad. The two ad versions are differently-sized cuts of the same raw footage, which are used as the two versions the advertiser used. We run the experiment for four days on over 50,000 users. The experimental variation in assignment to each ad format allows us to isolate the causal effect of these ad versions on different outcomes. It further enables us to isolate the substitution between ad and video consumption, as the assignment to the treatment exogenously shifts ad consumption.

To address our second challenge, we turn to the literature on the intersection of machine learning and causal inference. We use Causal Forests (Wager and Athey 2018) to estimate conditional average treatment effects (CATE) on both ad and video consumption outcomes. Finally, to address our third challenge, we develop a multi-objective personalization framework that takes CATE estimates for every outcome of interest (e.g., ad consumption and video consumption) as the inputs of a multi-objective optimization problem. We theoretically link the CATE estimates to both objectives, which helps us use the insights from the multi-objective optimization literature and design two algorithms called Greedy Front Elimination (GFE) and Parameter-Agnostic Scalarization (PAS). Both algorithms can navigate in the space of CATE estimates to identify policies that assign units to treatment conditions that balance the trade-off between ad and video consumption.

We first estimate the average treatment effect of the Skippable/Long ad version in our study on ad consumption relative to the Non-Skippable/Short ad. We find that exposure to the Skippable/Long ad results in 13.5 seconds higher ad consumption on average compared to the Non-Skippable/Short ad. We then define two binary outcomes based on ad consumption that relate to the completion of two important checkpoints: (1) reaching the second 15 of the ad[1], and (2) reaching the end of the ad. We find that, on average, a user is 1.9% more likely to reach the second 15 of the ad in the Skippable/Long condition compared to the Non-Skippable/Short condition. Given that users in the Skippable/Long condition have the option to skip the ad, this is a somewhat surprising finding that can serve as evidence that some users receive positive utility from consuming the ad, so they want to continue watching it beyond 15 seconds. When comparing the ad completion rate across the two groups, we find that the users in the Skippable/Long ad have a 32.2% lower chance of completing the ad compared to the users in the Non-Skippable/Short ad condition. This is expected because users in this condition need to watch a longer ad to reach the completion point while having the option to skip the ad.

Next, we focus on the video consumption as the outcome and find that assignment to the Skippable/Long ad version results in 9.5 percentage points lower video consumption compared to the Non-Skippable/Short ad. We include the control condition as a third condition and find that the average video consumption under the No-Ad condition is 15 and 24 percentage points higher than that under Non-Skippable/Short and Skippable/Long ad versions, respectively. Comparing the No-Ad condition with the two ad conditions, we find support for the ad avoidance mechanism, where the presence of an ad substantially reduces organic video consumption. Together with our results on ad consumption, we find empirical evidence for the presence of users with both positive and negative marginal utility of ad consumption.

We further explore the mechanism behind the findings and document that the effect of ad format on video consumption is fully explained by the impact of the ad format on ad consumption and whether the user starts the video. That is, the residual variation in video consumption when projected on ad consumption and video start indicator is not significantly different between the treatment groups. Finally, we test for the substitution between ad consumption and video consumption. We cannot regress video consumption on ad consumption simply because ad consumption is endogenous. However, since our treatment exogenously shifts ad consumption, we can use an instrumental variable research design. When we instrument ad consumption with the treatment assignment, we find a clear substitution pattern, where every 15 seconds more ad consumption results in approximately 13 seconds lower video consumption, on average.

Next, we document extensive heterogeneity in treatment effects on both ad and video consumption across time. We find that the magnitude of difference between the two ad versions

[1]We consider this binary outcome for two reasons. First, many platforms charge advertisers in the skippable condition if the user watches a certain amount of the ad, such as 15 seconds. Second, the second 15 checkpoint is defined for both ad conditions in our study.

is attenuated around the hours with the highest demand for the platform. We then use Causal Forests to estimate the heterogeneity in treatment effects more systematically. We show substantial variation in the distribution of CATE on both ad and video consumption. However, the distribution of CATE estimates for each outcome is largely unidirectional: all CATE estimates for ad consumption are positive, whereas nearly 97% of CATE estimates for video consumption are negative. This implies that only for 3.2% of all units, one treatment achieves higher ad and video consumption. Thus, even at the individual level, the platform faces a challenge in finding the right policy that increases both ad and video consumption.

Our multi-objective personalization framework aims to address the substitution between ad and video consumption. Intuitively, multi-objective algorithms identify units whose positive contribution to ad consumption outweighs their negative impact on video consumption and assign these units to the Skippable/Long ad version. We formulate the problem of multi-objective personalization in a generic manner, and design two algorithms for this task that directly link CATE estimates to personalized policies. We apply these algorithms to our data and empirically demonstrate their performance in both ad and video consumption using an Inverse Propensity Scoring (IPS) estimator (Horvitz and Thompson 1952). We prune the set of policies and drop those with a dominated performance in both objectives. The resulting curve under each algorithm is the Pareto frontier that mixes the assignment to Skippable/Long or Non-Skippable/Short ad formats to achieve reasonable performances in both objectives. To test how well these algorithms perform, we compare their performances with that of the random mixing of assignments. Our results reveal a large gap between the Pareto front generated by the algorithm and that of random policies, indicating that multi-objective personalization creates substantial value in this context.

To further quantify the value created by our multi-objective personalization algorithms, we compare their performance with that of single-objective personalization algorithms that only optimize with respect to one objective. In particular, we consider Single-Objective Ad Consumption (SOAC) and Single-Objective Video Consumption (SOVC) policies that find the personalized policy to optimize only ad consumption and video consumption, respectively. We document that compared to the single-objective personalized policy that only optimizes ad consumption, there is a multi-objective personalized policy that increases video consumption by 47% while only decreasing the ad consumption outcome by 13%. We further find larger gains when the platform wants to keep video consumption high. We show that compared to a single-objective personalized policy that only optimizes video consumption, there is a policy on the identified Pareto frontier that improves ad consumption by 61% while only reducing video consumption by 4%. Together, these findings show that multi-objective personalization can create value by substantially improving the performance in one dimension (e.g., video consumption) without hurting the performance in the other dimension (e.g., ad consumption).

In sum, our paper offers several contributions to the literature. Substantively, we document the interplay between ad and video consumption, two key performance metrics in the video advertising industry. In particular, we use a field experiment to causally establish the substitution pattern between these two consumption outcomes and measure their magnitude. We further provide empirical evidence for a less-supported claim that some users receive positive marginal utility from consuming ads. This finding has important implications for the optimal length and skippability of ads and emphasizes the importance of taking heterogeneity in treatment effect into account when making decisions. We document extensive heterogeneity in the effects of two ad formats in terms of ad and video consumption and show that although single-objective personalization fails to properly exploit this variation, multi-objective personalization can use this variation and achieve desired outcomes in both objectives. From a methodological standpoint, we propose a framework for multi-objective personalization and design two classes of non-parametric and parametric algorithms. We bring insights from the multi-objective optimization and apply them to the literature on confounding-robust policy evaluation. From a managerial perspective, our multi-objective personalization framework can be widely applied to other settings where there is a conflict in treatment effects on multiple desired outcomes. Our framework provides flexibility for policy-makers and managers who want to target a certain balance between outcomes by allowing them to evaluate the Pareto frontier a posteriori and select the policy.

2 Related Literature

Broadly, our paper relates to the advertising literature in marketing and economics. More specifically, our work examines the role of ad formatting in shaping user behavior. The question of how to optimally use different variations of the same ad has long been of interest to marketers (Schumann et al. 1990). With the reduction in the cost of experimentation, recent empirical studies have utilized field experiments to examine the impact of different formatting strategies, such as the well-known headline and ad size experiments for Bing's search ads (Kohavi and Thomke 2017), sponsorship disclosure (Sahni and Nair 2020), and spotlighting various features of the advertised product (Biswas 2020). In the context of video advertising, a few studies have investigated the impact features such as skippability (Pashkevich et al. 2012), ad length, and the presence of timer (Jeon et al. 2019) in the field or lab context. Closely related to our paper is Pashkevich et al. (2012), who run a field experiment on YouTube and show that the skippable ad format results in higher video consumption and user satisfaction compared to the non-skippable ad format. We add to this literature by providing experimental evidence that shows the opposite pattern. More importantly, we extend this literature by establishing the substitution pattern between ad and video consumption as the mechanism, which reconciles the findings of the two experiments under a more general framework.

Our paper also relates to the literature on the interplay between sponsored and organic content (Sun and Zhu 2013). With advancements in ad measurements, the literature on TV advertising has documented a phenomenon called "zapping", which is the practice of switching channels during commercial breaks (Zufryden et al. 1993, Danaher 1995, Siddarth and Chattopadhyay 1998). Since then, a series of papers have examined different aspects of ad avoidance by deriving the equilibrium properties in the market users are averse to ads (Anderson and Coate 2005, Dukes et al. 2022), quantifying the audience loss caused by ad avoidance the (Wilbur 2008), linking ad avoidance to sales (Bronnenberg et al. 2010, Deng and Mela 2018), and proposing market design solutions to account for audience externalities (Wilbur et al. 2013).[2] We contribute to this stream of work by causally identifying a substitution pattern between sponsored and organic content consumption. Importantly, we show that the channel for this substitution is not only ad avoidance but also the fact that some users have a positive marginal utility of ad consumption. We further 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).

On the methodological front, our paper relates to the literature on multi-objective optimization (Marler and Arora 2004). This literature has proposed a series of algorithms to deal with multi-objective optimization problems, ranging from scalarization techniques (Miettinen and Mäkelä 2002) to genetic algorithms (Deb et al. 2002). More closely related to our paper is the stream of literature that considers discrete policies that map the covariate vector to a specific treatment condition, such as the literature on multi-objective contextual bandits (Tekin and Turğay 2018, Turgay et al. 2018, Wang et al. 2023) and multi-objective reinforcement learning (Roijers et al. 2013, Van Moffaert and Nowé 2014, Abdolmaleki et al. 2020). Our paper extends this stream of work by bringing insights from the causal inference literature and designing algorithms that directly incorporate Conditional Average Treatment Effect (CATE) estimates when developing personalized policies.

Finally, 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 2013, 2019). Recent methodological developments in this literature have brought a causal lens to machine learning algorithms that have been traditionally used for personalization tasks (Athey and Imbens 2016, Shalit et al. 2017, Wager and Athey 2018, Nie and Wager 2021). Applied papers in this domain have documented the gains from personalization in a variety of domains, such as incentives in churn management problems (Ascarza 2018), promotional offers in

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

retail settings (Simester et al. 2020a,b), allocation and sequencing of mobile in-app advertising (Rafieian and Yoganarasimhan 2021, Rafieian 2022), length of free trial in software as service industry (Yoganarasimhan et al. 2022), and product versioning in music streaming platforms (Goli et al. 2022). The key insight in this series of work is that having a fine-grained set of pre-treatment variables helps differentiate between users, thereby creating value by assigning users to the right policy. We extend this literature by proposing a multi-objective personalization framework that allows firms to identify policies that generate considerable gains in many dimensions by exploiting the variation in CATE estimates across different outcomes. In particular, we show that even in a context where single-objective personalized policies offer limited differentiation, a multi-objective personalization approach can create substantial value by differentiating between users based on the magnitudes of treatment effects and substitutability between outcomes at the individual level. Our generic and flexible framework makes it applicable to many marketing and non-marketing problems where the decision-maker needs to optimize more than one objective.

3 Experiment

In this section, we describe our experimental context. We start with the setting of our study. We then describe the exact experiment we deliver. Next, we describe our data and present some important summary statistics.

3.1 Application Setting

The application setting of our study is the video advertising industry. We partner with the company `vdo.ai`, which is based in India and the US and provides video services to publishers worldwide. Since its launch, `vdo.ai` has attracted many large and medium-sized media publishers who use the company's technology to serve video content and video ads on their platforms.

As a form of monetization, `vdo.ai` places video ads at different parts of the organic video. Video ads are generally of three types: (1) pre-roll ads that are placed prior to the start of the video, (2) mid-roll ads that are placed in the middle of the video, and (3) post-roll ads that are placed after the content video has finished playing. Figure 1 visualizes these different types of ads. In our experiment, we only focus on the pre-roll ads that are shown before the organic content starts.

The platform uses two different inventories of 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. 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

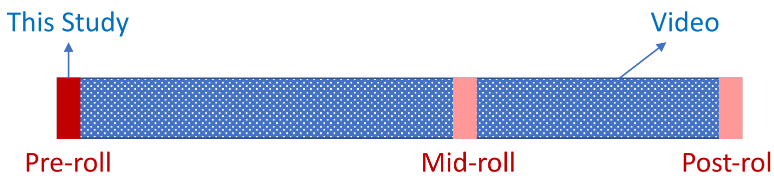


Figure 1. A visualization of different types of video ads based on their places within the organic video content.

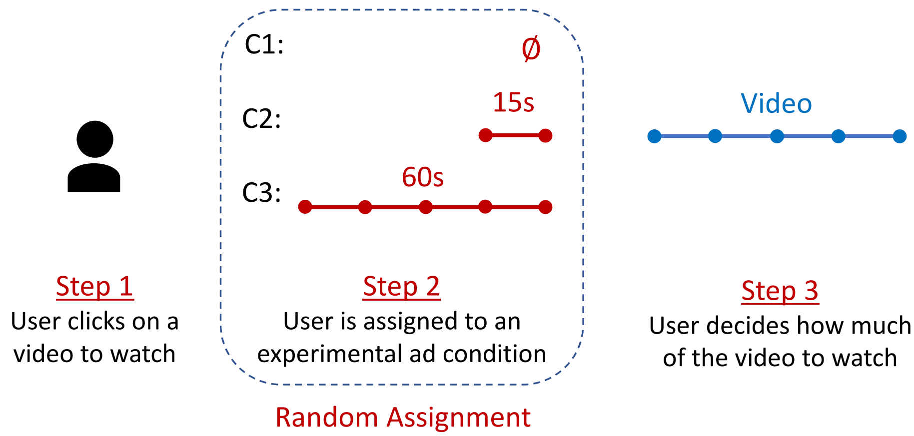


Figure 2. A visualization of our experiments.

targeting process.

3.2 Experiment Design

We design a fully randomized experiment at `vdo.ai`, where the pre-roll impressions are assigned to three experimental conditions: (1) No-Ad condition, where the user does not need to watch an ad to start consuming the organic video content, (2) Non-Skippable/Short ad condition, where the ad shown is a 15-second long ad of the `boAt`'s Watch Xtend product that is no skippable, and (3) Skippable/Long ad condition, where the ad shown is a 60-second long ad of the same `boAt` product that is skippable after five seconds. Figure 2 shows a schema of our experiment and the treatment conditions. The points show the trackers placed to find whether the user has reached a certain point in the ad and video. This means that we can record whether the user has reached the midpoint of the Skippable/Long ad (i.e., second 30) or the third quarter of the organic video (75%). Although we can control for the exact ad shown, the organic video content is chosen by the users, so they can be different videos with different lengths. However, since we ran a randomized experiment, the organic video does not affect the treatment condition. Thus, the distribution of videos is the same across treatment conditions.

For the Non-Skippable/Short and Skippable/Long ad conditions, we use two versions of an advertisement for the same `boAt` product. Figure A1 in Web Appendix A shows a snapshot of different parts of both ad versions. The two ads are just short and long cuts of the same raw footage. In that sense, these ads are the short and long ad versions launched by an actual advertiser. However, we acknowledge the limitation that ads can differ on dimensions other than skippability and length. Thus, our analysis does not explicitly focus on isolating the effect of these two features alone. Rather, we adopt the perspective of a platform that wants to choose between the two ad formats.

In total, we split the impressions randomly across treatment conditions with different weights, such that the No-Ad condition is used for 10% of all impressions, and either one of the ad conditions is shown in 45% of impressions each. We ran the experiment for four days, from July 19–22 in 2022. Our data comes predominantly from India. In the next section, we share more about our data.

3.3 Data

Each observation in our data refers to a unique video-watching experience, which we define as a session. Each session originates from a user's request to watch a video and is assigned to one of the three experimental conditions defined earlier: (1) No-Ad, (2) Non-Skippable/Short, and (3) Skippable/Long. In total, there are 59,692 sessions in our data that are generated by 57,343 unique users. This means that we have more than one session for many users. Because randomization has been implemented at the session level, one user may be assigned to multiple treatment conditions in different sessions. As such, we only focus on the first session for each user for our analysis to avoid any interference bias.

3.3.1 Variables

For each session in our data, we observe the following sets of variables:

- Pre-treatment variables: This set contains the user's *IP Address*, *Time of Day*, *Date*, *City*, *Country*, and the *Operating System (OS)* of the device that the user is using.
- Treatment variable: Each session is randomly assigned to a treatment condition, where the treatment conditions are presented as follows: (1) *No-Ad*, (2) *Non-Skippable/Short*, and (3) *Skippable/Long*.
- Outcome variables: We collect a rich set of post-treatment variables or outcomes both on the ad performance and video engagement metrics. As shown in Figure 2, we place trackers at different points in the sessions that indicate whether the user has reached those points. For both ads, the trackers are placed every 15 seconds[3], and for the video, these trackers are placed

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

for every quarter (25%) of the video.[4] In addition to the trackers shown in Figure 2, 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 (CTR) is relatively low.[5]

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. In particular, we have detailed information on the consumption of sponsored and organic content in each session.

Such detailed tracking also allows us to identify whether a user faces technical issues or uses an ad blocker. In particular, if the tracker at the beginning of both the ad and video returns null values, we assume that the user had technical issues, such as a network problem. Similarly, if the tracker 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.[6] 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 and 56,662 unique users to work with. Since we only use the first session for each user, our final sample has a total of 56,662 sessions to study.

3.3.2 Summary Statistics

We now present some basic summary statistics of the data. We start with the pre-treatment variables, which are all categorical variables. We find the top three subcategories with the highest number of observations for each variable in our data. We present this information about each variable along with the total number of subcategories in Table 1. As shown in this table, the hours with the highest traffic are 6–8 AM MST, which would be 5:30-7:30 PM in India, where most of the traffic comes from. The experiment was run from July 19 through July 22, and the last two days had the highest traffic.

As indicated in Table 1, there are a total of 956 cities in our data. However, over half of the observations are from Mumbai. It is worth noting that there are many cities with only one observation in our data. Next, we find that the vast majority (99.56%) of all observations occur in India. The statistics on our final pre-treatment variable show that Android OS is the most common OS in our data, with around 80% of the total traffic. In Web Appendix B, we perform

[4] 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 long video. This is a limitation of our analysis. However, since we randomize the treatment, the video lengths would not significantly differ across groups.

[5] 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).

[6] It is clear that we cannot identify ad blockers for users in the No-Ad condition. Since our main analysis concerns the difference between two ad formats, this does not cause a problem in our main analysis.

Variable	Number of subcategories	Top three subcategories and their shares		
		1^{st}	2^{nd}	3^{rd}
Hour of Day	24	7AM MST (7.59%)	8AM MST (6.75%)	6AM MST (6.67%)
Date	4	07/22/2022 (38.51%)	07/21/2022 (38.27%)	07/20/2022 (14.48%)
City	956	Mumbai (51.21%)	Delhi (7.52%)	Hyderabad (6.85%)
Country	12	India (99.56%)	United States (0.38%)	Spain (0.02%)
Operating System	6	Android (79.50%)	Windows (12.90%)	iOS (5.59%)

Table 1. Summary statistics of the pre-treatment variables.

extensive randomization checks on the distribution of pre-treatment variables to ensure that the randomization has been implemented correctly in our study.

4 Empirical Analysis of the Experiment

In this section, we present an empirical analysis of the experiment described in the previous section. We start with the analysis of the average treatment effect in §4.1. Next, in §4.2, we explore the mechanism behind our findings. Finally, in §4.3, we extend our analysis by estimating heterogeneous treatment effects using the pre-treatment covariates.

4.1 Average Treatment Effect

Before we proceed with the analysis, we need to define some notation. Let i denote each observation in our data, and X and W denote the pre-treatment covariates and the treatment variable, respectively. Since we are interested in the difference between the ad formats, we define W as a binary variable with $W = 1$ and $W = 0$ referring to Skippable/Long and Non-Skippable/Short ad formats, respectively. We use the control condition (No-Ad) for analysis when needed.

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

For inference, we use the common assumptions in the causal inference literature: (1) overlap, (2) unconfoundedness, and (3) the stable unit treatment value assumption (SUTVA). The first two are satisfied by design since we have a randomized controlled trial. SUTVA is also reasonable because there is no interaction between users, and the treatment received by all users in the same treatment condition is identical (i.e., no multiple versions of the treatment). Under these assumptions, we know that the average treatment effect is the difference in group averages (Neyman 1923). We use this fact for our main analysis.

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

Table 2. Description of outcome variables.

Finally, we stress that because our treatments differ in multiple dimensions, we can attribute the identified causal effect to any differences between the two ad formats. To that end, our estimated treatment effects are a composite of three factors that differ between ads: (1) skippability, (2) length, and (3) other ad-specific differences in content. Although the two ads are chosen to minimize the role of other ad-specific differences, it is not possible to entirely rule out this channel. Thus, we must be wary of this effect decomposition when interpreting the mechanism behind our treatment effects.

4.1.1 Average Treatment Effect on Ad Consumption

In this section, we examine the effect of using a Skippable/Long ad format on a series of ad-related outcomes, compared to the Non-Skippable/Short ad format. From a theoretical perspective, it is not clear which ad format leads to higher ad consumption, which is measured by the time spent viewing the ad content. On the one hand, a longer ad has an inherent advantage as it can be consumed for a longer time. On the other hand, the ability to skip the long ad after 5 seconds may result in lower consumption of the longer ad. For example, if most people skip the longer ad after 5 seconds, we expect the average ad consumption to be lower for the Skippable/Long ad compared to the Non-Skippable/Short ad. Therefore, the effect of ad format on ad consumption is an empirical question. The first row in Table 3 provides an answer to this empirical question in the context of our study: we find that the average consumption of the Skippable/Long ad is

Outcome	Mean of Treatment A (Non-Skippable/Short)	Mean of Treatment B (Skippable/Long)	Mean Difference $B - A$ Estimate	p-value
Ad Consumption ($\times 15s$)	0.53526	1.43567	0.90041	< 0.001
Second 15 Complete	0.53526	0.55413	0.01887	< 0.001
Ad Complete	0.53526	0.21338	-0.32188	< 0.001
Ad Click	0.00125	0.00128	0.00003	0.926

Table 3. Average treatment effect for the ad-related outcomes. The number of observations is 51,423 for all models.

significantly higher than that of Non-Skippable/Short ad, with the average treatment effect being $0.90 \times 15 = 13.50$ seconds, which is approximately equal to the length of the short ad in our study. Thus, users spend greater time viewing the Skippable/Long ad in our study.

To further pin down the extent to which this higher ad consumption in the Skippable/Long ad condition is driven by the inherent difference in the length of the ad, we examine which ad format has a higher likelihood of completing the first 15 seconds of the ad. As such, we use the binary outcome *Second 15 Complete*, as defined in Table 2. With the same length of consumption, we can better examine the role of the skippability option, as users in only one condition can skip the ad. The conventional wisdom is that the Skippable/Long ad format will have a lower likelihood of consuming 15 seconds of the ad. Surprisingly, we find the opposite pattern in the second row of Table 3: despite the presence of the skip option in the Skippable/Long condition, the completion rate of the first 15 seconds is significantly higher in Skippable/Long condition compared to the Non-Skippable/Short condition. We find that only 1.9% of users in the Skippable/Long ad condition skip the ad before the Second 15 checkpoint to start watching the video, and 42.7% of them completely avoid the ad. Assuming that the utility of consuming the first 15 seconds of both ads is the same for each user, one likely explanation for this finding is that the skippability option reduces immediate ad avoidance, thereby allowing more users in the Skippable/Long condition to tune in and consume the ad.[7] More generally, this finding indicates that many users receive a positive utility from consuming the Skippable/Long ad.

Next, to test the extent to which users exposed to the Skippable/Long version continue watching the ad, we compare *Ad Complete* for two ads. Theoretically, we expect a higher ad completion under Non-Skippable/Short ad because shorter ads are easier to complete, and the inability to skip forces users in this condition to complete the ad in order to watch the organic video content. As expected, the third row in Table 3 demonstrates this point: 53.5% of Non-Skippable/Short

[7] We acknowledge that this assumption may not hold because ads differ in their first 15 seconds. However, using more detailed data on the Non-Skippable/Short ad, we see that over 20% of users close the session before reaching its first quarter, which happens after only 3.75 seconds. That is, the lack of a skip button likely results in users immediately closing the tab.

ads are completed, whereas only 21.3% of Skippable/Long ads are completed. Lastly, we focus on users' click decision on ads as the final outcome (*Ad Click*). As discussed earlier, the objective of the `boAt` ad campaign in our study is to generate more awareness. As such, although the ad is clickable, it is not a performance ad with a clear call for action. Click only takes users to the product's website for more information. The fourth row of Table 3 compares the performance of the two conditions in terms of Ad Click. Both ads generate around 0.1% click-through rate (CTR), and the difference is not statistically significant.

In summary, our analysis reveals that the Skippable/Long ad format in our study results in higher ad consumption than the Non-Skippable/Short ad format. This is an important result from a revealed preference point-of-view, as users in the Skippable/Long condition have the choice to skip the ad at any point five seconds into consuming the ad. Together with a high rate of abandoning the session, our results indicate that the consumption utility of the ad in our context is largely heterogeneous, which motivates our study's focus on personalization.

4.1.2 Average Treatment Effect for Video Consumption

While ad-related outcomes are important to advertisers, the video advertising platform and the content creators also care about video consumption. Since we work with pre-roll ads, we expect the ad format to affect both of these outcomes. Prior research provides empirical evidence for ad avoidance, and the negative effect ads have on the consumption of organic content (Wilbur 2008, Wilbur et al. 2013, Goli et al. 2022). Since skippability facilitates ad avoidance without abandoning the session, a general expectation is that users will consume more organic content when they can skip ads. Prior studies on ad skippability have corroborated this insight by showing a higher rate of organic content consumption when users are exposed to skippable ads (Pashkevich et al. 2012). In this section, we estimate the average treatment effect on video consumption that measures how many quarters of the video each user has consumed. Interestingly, our results in the first line of Table 4 show the opposite pattern: users in the Skippable/Long condition consume 0.38 quarters less than those in the Non-Skippable/Short condition. This is equivalent to a 9.5 percentage point difference in the video consumed.

We break down video consumption into four binary variables for each quarter of video that is reached. As shown in Table 4, all quarters are more likely to be reached in the Non-Skippable/Short condition than the Skippable/Long condition. Specifically, we find that the difference in video consumption starts from the beginning of the video. As shown in Table 4, we find that 44.5% of users in the Non-Skippable/Short condition started watching the video, which is lower than 53.5% who completed the ad, indicating that there is some dropout in the transition from ad to video. For the Skippable/Long ad condition, the user can start watching the video without necessarily completing the ads, as the ad is skippable in this condition. We find that 20.7% of users in this

Outcome	Mean of Treatment A (Non-Skippable/Short)	Mean of Treatment B (Skippable/Long)	Mean Difference $B - A$ Estimate	p-value
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

Table 4. Average treatment effect for the ad-related outcomes. The number of observations is 51,423 for all models.

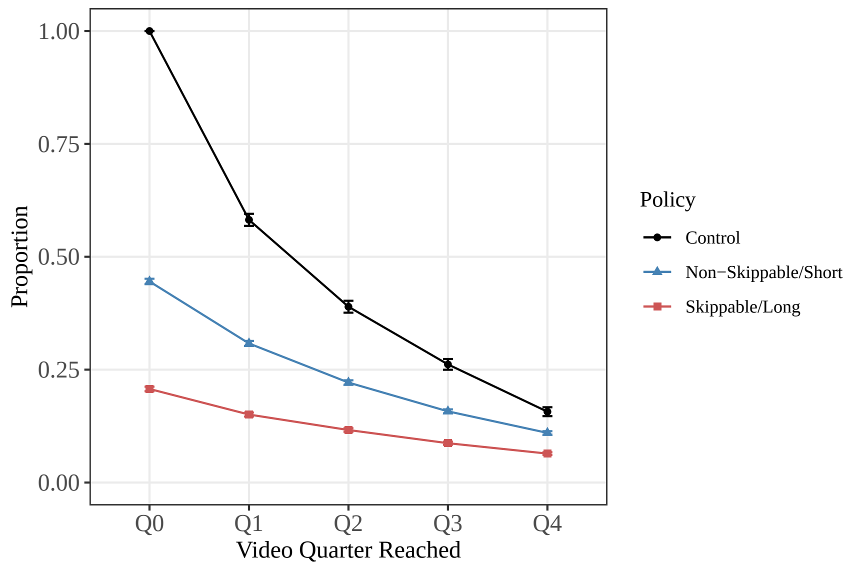


Figure 3. Proportion of users reached each quarter of the video across three policies. Error bars are 95% confidence intervals for proportions.

condition have started watching the ad, 14.1% of whom have reached there by skipping the ad.

We further compare the video consumption measures of these two ad formats with the control condition where there is no ad. Figure 3, we show the fraction of users who have reached each quarter across the three treatment conditions. The blue and red lines show the proportions for Non-Skippable/Short and Skippable/Long ad formats, respectively, and provide a visual presentation of the information in Table 4. The black line shows the video consumption in a condition where no ad is present. As a result, the video start rate is 1, and the fraction of surviving users decreases over the course of the video. The fraction of users who reached each quarter of the video is significantly higher for the control condition than both the conditions with a pre-roll ad before the video. This finding highlights the role of ad avoidance in our study, as both conditions with an ad result in substantially lower video consumption.

Another interesting pattern that emerges from Figure 3 is the difference in the rates at which the fraction of surviving users declines across different policies. This figure shows that the No-Ad condition has the steepest negative slope, whereas the Skippable/Long ad condition has the flattest negative slope. It is worth emphasizing that this is not a causal effect, and users who are consuming the video are different in the two conditions. For example, the sample of users who are present at the beginning of the video in the Skippable/Long condition (approximately 20% of users in this treatment arm) is likely a selected sample of users who are more interested in the video, thereby creating a higher conditional survival rate.

In summary, we show that in our study, the Non-Skippable/Short ad leads to a higher video consumption than the Skippable/Long ad. Although we acknowledge that the treatment effects can vary depending on the advertised product, we highlight an important point that the impact of these ad formats highly depends on the ad itself and how users consume it. Thus, in the absence of a generalizable impact of these ad formats, it is crucial for platforms to better understand the mechanism behind driving ad and video consumption.

4.2 Mechanism Analysis

In this section, we discuss the mechanism behind our findings. We first want to understand what explains the treatment effect on the video consumption outcome. The pre-roll ad can affect the video consumption outcome in two ways. First, it can affect users' intention to even start watching the video. This pattern is clearly shown in Figure 3: more than half of the users in either ad condition do not start the video. Second, ads are video contents themselves that users consume. As a result, ad consumption can affect video consumption at an intensive margin. Together, we hypothesize that the treatment effect on video consumption is fully explained by the treatment effect on two outcomes: (1) *Ad Consumption*, and (2) *Video Start*.

To test this hypothesis, we can examine whether the residual variation in video consumption is significantly different across Skippable/Long and Non-Skippable/Short ad conditions when accounting for both Ad Consumption and Video Start. To do so, we first regress Video Consumption on Ad Consumption and Video Start to obtain the residuals. We then regress the residual variation in video consumption on our treatment variable. We present the result of this practice in column 1 of Table 5. Our results indicate that there is no significant difference between the ad versions in their residual video consumption. This finding suggests that the treatment effect on Ad Consumption and Video Start fully explains the treatment effect on video consumption.

We further investigate the relationship between ad and video consumption. One 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

	Dependent variable		
	(1) Residual Video Consumption	(2) Video Consumption	(3) Video Consumption
Treatment	−0.0045		
	(0.0080)		
Ad Consumption		0.3538***	−0.4209***
		(0.0042)	(0.0150)
Instruments	None	None	Treatment
Weak Instruments			7567***
No. of Obs.	51,423	51,423	51,423
Note:			*p<0.05; **p<0.01; ***p<0.001

Table 5. Regression results for mechanism analysis. Numbers reported in parenthesis are standard errors of estimates.

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 can use a two-stage least square (2SLS) estimator where in the first stage, we regress the endogenous variable Ad Consumption on our instrument variable Treatment by estimating the following model:

$$AdConsumption_i = \alpha_0 + \alpha_1 Treatment_i + \nu_i \tag{1}$$

We then use the estimates of this model to estimate Ad Consumption and obtain $\widehat{AdConsumption}_i$ to plug it in the second stage model as follows:

$$VideoConsumption_i = \beta_0 + \beta_1 \widehat{AdConsumption}_i + \epsilon_i, \tag{2}$$

where β_1 is our coefficient of interest that determines the degree of the substitution between ad and video consumption. In columns 2 and 3 of Table 5, we present the results from both plain and Instrumental Variable regressions. Although the results of column 2 show a positive association between ad 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 3 of Table 5. 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 2 minutes or 120 seconds. Using a back-of-the-envelope calculation, we find that every 15-second increase in ad consumption decreases video consumption by $120 \times 0.1052 = 12.62s$, on average.

Establishing the substitution between ad and video consumption is important as it provides a more generalizable understanding of why skippability does not always lead to higher video consumption. From a platform perspective, this substitution pattern highlights an inherent trade-off in optimizing ad and video consumption. That is, at the aggregate level, strategies that increase ad consumption come at the expense of video consumption. Therefore, platforms need more advanced tools to achieve the right balance.

4.3 Heterogeneity in Treatment Effects

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 (i.e., ad and video consumption). To find the right balance, it is crucial to understand whether there is any heterogeneity in treatment effects on both outcomes, and if so, where this heterogeneity comes from. To characterize this heterogeneity, we turn to a new estimand: Conditional Average Treatment Effect (CATE). The definition of CATE is the same as ATE, given a specific value of covariates. For the set of potential outcomes $Y_i(1)$ and $Y_i(0)$, we denote CATE at $X = x$ by $\tau(x)$ and define it as follows:

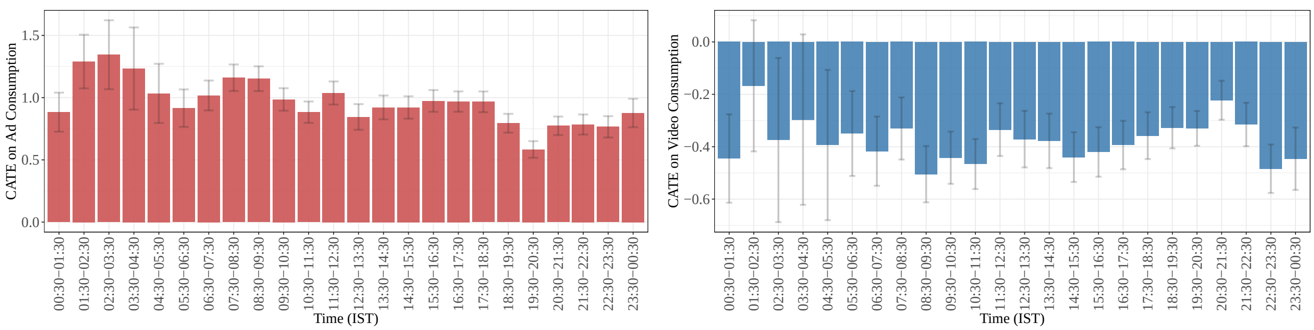
$$\tau(x) = \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = x]. \tag{3}$$

In this section, we work with this estimand and first explore the heterogeneity across time in §4.3.1, and then present a more systematic approach to estimate heterogeneity across the entire covariate space in §4.3.2.

4.3.1 Heterogeneity in Treatment Effects Across Time

From a theoretical perspective, we expect the time of the day to influence users' consumption utility from watching sponsored and organic content. As such, we estimate the treatment effects at any specific hour of the day. Because we have an experiment, we know that treatments are properly randomized at any given point in time. As a result, we can use a simple mean difference estimator. We only focus on our sample in India as it constitutes 99.5% of all observations, and we modify the time zone from MST to IST for a more meaningful interpretation.

We estimate CATE on both outcomes for every hour of the day and present the results in Figure 4. Figure 4a shows the hour-level heterogeneity in treatment effects on ad consumption, whereas Figure 4b shows this heterogeneity in treatment effects on video consumption. Both figures reveal interesting patterns. We find that there is considerable heterogeneity in treatment effects across time. We further notice a gradual decline in the magnitude of treatment effects on both outcomes after 5:30 PM till 8:30 PM. This indicates that users' consumption is overall less sensitive to the ad format. One possible explanation for this pattern is that it is during the more focused leisure time of the users, so they have a clearer goal of what organic video to consume. After 9:30 PM, we observe that the CATE on ad consumption increases, which is likely the main driver behind the



(a) Outcome: Ad Consumption (×15 Seconds) (b) Outcome: Video Consumption (in Quarters)

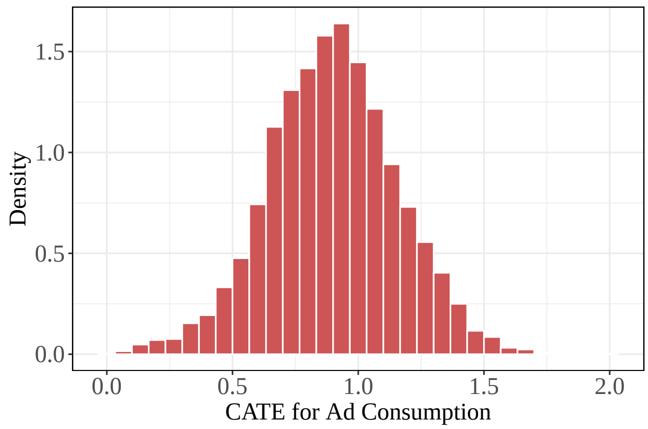
Figure 4. Heterogeneity in treatment effects on both Ad Consumption and Video Consumption across time of day. Times are presented in Indian Standard Time (IST). Error bars are 95% confidence intervals around treatment effects.

increase in the magnitude of the CATE for video consumption. However, this substitution pattern is not as strong during midnight hours. That is, users have higher ad consumption under the Skippable/Long ad format, but it does not result in lower video consumption. Finally, we notice that the substitution pattern becomes stronger during users' work hours, with a highly positive treatment effect on ad consumption and a highly negative treatment effect on video consumption. This is expected as users have more limited time during work hours.

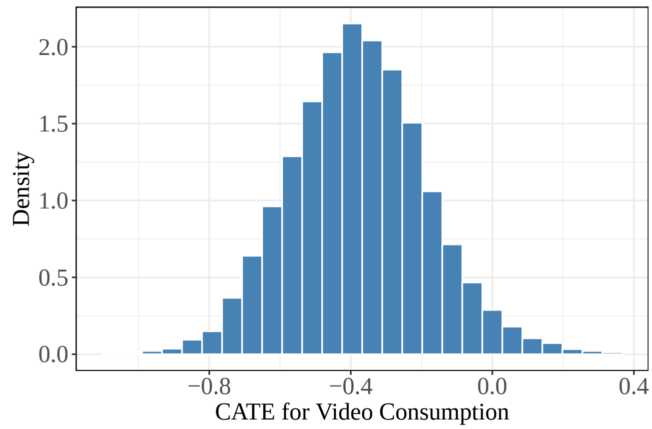
Overall, the patterns in Figure 4 show heterogeneity in treatment effects across times of the day. More importantly, we find some variability in the extent of substitution at different time periods. We later use this intuition for our multi-objective personalization.

4.3.2 Heterogeneity in Treatment Effects Across All Covariates

Although the patterns in Figure 4 shows extensive heterogeneity in treatment effects across times of day, the substitution pattern between ad and video consumption is still quite apparent: for all hours of the day, the treatment effect on ad consumption and video consumption have opposite signs. This motivates us to capture the heterogeneity in treatment effects more systematically at the individual level. If the positive average treatment effect on ad consumption and the negative average treatment effect on video consumption come from separate portions of our data, the solution is clear for the platform. For example, suppose that there are two groups of users $\mathcal{I}_a$ and $\mathcal{I}_v$ such that $\mathcal{I}_a \cap \mathcal{I}_v = \varnothing$, where users in $\mathcal{I}_a$ have a positive CATE on ad consumption and a positive CATE on video consumption, whereas users in $\mathcal{I}_v$ have a negative CATE on 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.



(a) Outcome: Ad Consumption ($\times 15s$)



(b) Outcome: Video Consumption (in Quarters)

Figure 5. The distribution of CATE estimates for both Ad Consumption and Video Consumption as outcome variables.

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

We present the histogram of our CATE estimate for both ad and video consumption outcomes. Figure 5a shows how CATE for the Ad Consumption outcome varies across individuals. As shown in this figure, although there is extensive variation in the CATE estimates, the sign for all units remains positive. This indicates that the Skippable/Long ad format results in greater ad consumption compared to Non-Skippable/Short format for all individuals in our data. We find that the CATE estimates are significant for 97.43% of all users. Overall, our results indicate that there is no user whose CATE estimate is negative and significant. Thus, if we use this sole objective for developing a personalized policy, the resulting policy will be a uniform Skippable/Long ad condition for everyone.

We then move on to CATE estimates for Video Consumption as our video-related outcome and visualize the distribution of CATE estimates in Figure 5b. As shown in this figure, although the vast majority of CATE estimates are negative, there is a small 3.15% of users with positive CATE estimates. When considering the significance of these estimates, we find that only three users (less than 0.01% of the total users) have positive and significant CATE estimates, whereas 55.13% of those have negative and significant CATE estimates. Together, the optimal personalized policy

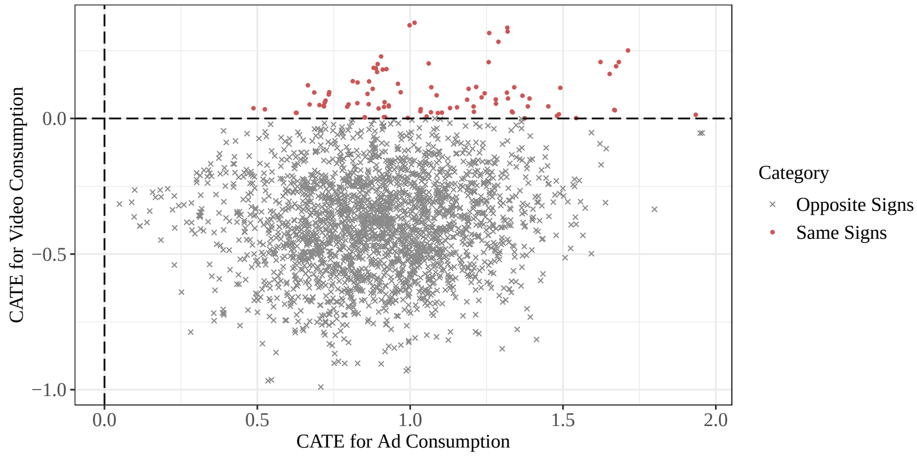


Figure 6. Scatter plot of CATE estimates for Video Consumption (in quarters) against CATE estimates for Ad Consumption ($\times 15s$).

with respect to Video Consumption as the objective is almost the same as a uniform policy where all users are assigned to a Non-Skippable/Short ad.

Although we find substantial variation in the heterogeneous treatment effects on both ad and video consumption outcomes, the substitution pattern persists even at the individual level. To better understand the substitution pattern at the individual level, we plot the CATE on video consumption against the CATE for ad consumption and present the resulting scatter plot for a random sample of our observations in Figure 6. The first pattern that emerges from this figure is that only for a small portion of units do we have the same sign for CATE on both outcomes. These points (shown in red in Figure 6) account for 3.15% of all units in our data.

Finally, we ask a broader question: To what extent are CATE estimates for these two outcomes in conflict at the individual level? Since we want higher CATE estimates for both outcomes at the individual level, we want a more positive correlation between these CATE estimates. On the other hand, a negative correlation between these CATE estimates indicates that a higher CATE for one outcome is associated with a lower CATE for another outcome, thereby making the multi-objective solution more challenging. As shown in Figure 6, there is a weak positive correlation between CATE estimates for both outcomes (correlation $= 0.13$). Although the positive association between these CATE estimates is not strong, it is still promising as it suggests that CATE estimates move in the same direction, on average. Intuitively, points that contribute most to higher ad consumption have a more positive (or less negative) impact on video consumption. Thus, the pattern in Figure 6 suggests that multi-objective personalization can be useful for the platform that wants to simultaneously achieve higher ad and video consumption. 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 Multi-Objective Personalization

Motivated by the challenge presented in the previous section, our goal in this section is to perform multi-objective personalization. Since many applied problems in marketing involve working with multiple objectives, we first present a generic case of the problem and propose solutions that are not specific to our setting. We then take these solutions to data and examine how they work in our context. In the following sections, we first formally define our problem in §5.1. Next, in §5.2, we present the two algorithms we use for the problem at hand. In particular, we propose a greedy algorithm for our task at hand called *Greedy Front Elimination*. We then draw upon the insights from the multi-objective optimization literature on scalarization and use an algorithm called *Parameter-Agnostic Scalarization*. In §5.3, we present our strategy for counterfactual policy evaluation. Finally, in §5.4, we use our data to evaluate the performance of different multi-objective personalization algorithms and examine whether there is any gain from doing so in our context.

5.1 Problem Definition

At a high level, multi-objective personalization entails developing a personalized policy that performs well in terms of multiple objectives. To characterize this problem, we need to first define what we mean by a *personalized policy* and *performance*. We follow the conventional definitions in this literature as laid out in Rafieian and Yoganarasimhan (2023) to define a personalized policy. As before, 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, in a binary treatment context, we have $\mathcal{W} = \{0, 1\}$. Now, we can define a policy π as follows:

Definition 1. *A policy $\pi : \mathcal{X} \to \mathcal{W}$ is a mapping from the covariate space to the treatment space. That is, for any vector of covariates $X_i \in \mathcal{X}$, the treatment prescribed by the policy is $\pi(X_i)$.*

With this definition of a policy, finding a policy is a search over the space of $|\mathcal{W}|^N$ policies, where N is the number of observations.[8] To perform this search effectively, we need a performance measure tied to our multiple objectives. For example, in our context, we want to know how each policy performs in terms of ad and video consumption. We define these performance measures as follows:

Definition 2. *For each outcome Y, we define the performance of the policy in terms of that outcome as a mapping $\rho_Y : \Pi \to \mathbb{R}$, where Π is the space of all possible policies. This indicates*

[8] It is worth noting that Definition 1 only allows for deterministic policies. A more general alternative that allows for probabilistic policies is to define $\pi(x|w)$ as a conditional distribution, where for each value of the covariate space, each treatment has a probability of being prescribed. Our proposed algorithms can easily be extended to a probabilistic setting.

that for a policy π, the performance in terms of outcome Y is characterized by $\rho_Y(\pi)$. We can formally define this term as follows:

$$\rho_Y(\pi) = \mathbb{E}\left[Y_i\left(\pi(X_i)\right)\right], \tag{4}$$

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

There are many ways to characterize this objective using data. For example, one could use CATE estimates that measure the impact of the treatment for each user. We keep this objective abstract for our problem definition to allow for different forms of estimating the expected outcome. However, we later link it to the data using our CATE estimates when describing different algorithms.

In a multi-objective personalization problem, many different performance metrics can come into play. As a result, comparing two policies is more challenging in a multi-objective case compared to a single-objective case. For example, in the context of our problem, we can consider two objectives $\rho_A(\pi)$ and $\rho_V(\pi)$ that are defined for the Ad Consumption and Video Consumption outcomes, respectively. If there are two policies π_1 and π_2 such that $\rho_A(\pi_1) > \rho_A(\pi_2)$ and $\rho_V(\pi_1) < \rho_V(\pi_2)$, it is not clear which one the platform must choose. However, if there are two policies π_1 and π_2 such that $\rho_A(\pi_1) > \rho_A(\pi_2)$ and $\rho_V(\pi_1) > \rho_V(\pi_2)$, we can conclude that policy π_2 is dominated by π_1 with respect to both objectives. This type of comparison immediately brings us to the notion of Pareto optimality, where the Pareto frontier of the policy space is the set of policies that are non-dominated by any other policies. To this end, we define the main goal of multi-objective personalization as follows:

Definition 3. *Suppose that there is a policy-maker who wants to optimize multiple outcomes Y_1, Y_2, ..., Y_K. Our goal is to find a set of policies Π_f that are Pareto optimal in terms of objectives $\rho_{Y_1}(\pi)$, $\rho_{Y_2}(\pi)$, ..., $\rho_{Y_K}(\pi)$. That is, for each $\pi \in \Pi_f$, there is not other policy π' in the space of policies such that we have $\rho_{Y_j}(\pi') > \rho_{Y_j}(\pi)$, for every j.*

The literature on multi-objective optimization offers many solutions to this problem, given the setting. Much of this literature focuses on the problem with a set of continuous control variables set by the policy-maker that are linked to multiple notions of reward or objective (Marler and Arora 2004). For example, a driver can set continuous variables speed and total passenger weight to optimize the travel time and fuel cost. Multi-objective personalization problem involves finding a complex discrete policy that performs well on multiple objectives. As such, our problem is more closely related to the literature on multi-objective contextual bandits (Tekin and Turğay 2018, Turgay et al. 2018, Wang et al. 2023) and multi-objective reinforcement learning (Roijers et al. 2013,

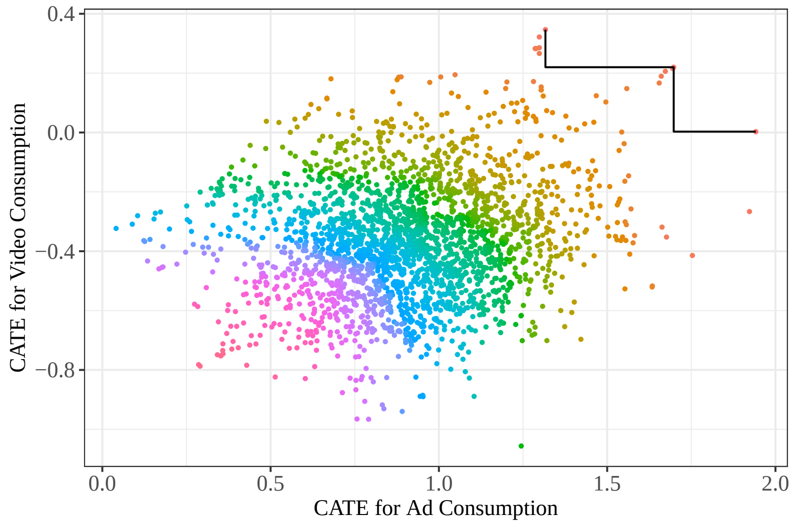


Figure 7. Pareto fronts of the sample of our data in terms of CATEs for Ad Consumption and Video Consumption.

Van Moffaert and Nowé 2014, Abdolmaleki et al. 2020). Many of the proposed solutions suggest some form of scalarization where the objectives are weighted and transformed into a single objective, or a constrained optimization where we optimize one objective while satisfying constraints on the other objectives. We draw upon the common approaches used in this literature and design algorithms that directly incorporate CATE estimates into the multi-objective personalization problem. We discuss these algorithms in the next section.

5.2 Algorithms for Multi-Objective Personalization

As discussed earlier in Definition 3, our goal is to design algorithms that find a set of Pareto optimal policies. An important task is to link CATE estimates for each objective to the policy performance under that objective. In this section, we first propose a simple greedy algorithm and then present the conventional scalarization solution that has been widely used in the literature.

5.2.1 Greedy Front Elimination Algorithm

We start by proposing a greedy algorithm for multi-objective personalization. Since our ultimate goal is to find the Pareto frontier of the policy performance in $\mathbb{R}^K$, we turn to the Pareto frontier of the CATE space in $\mathbb{R}^K$, which is closely related to the policy performance. The Pareto frontier of the CATE space are the set of observations whose CATEs are non-dominated by any other observation. Figure 7 shows the Pareto frontier of the CATE on Ad Consumption and Video Consumption for a random sample of our data as a step function in the top right of the graph. Intuitively, we expect the Skippable/Long ad format to be more valuable for these points than the other Pareto-dominated points.

We further define the notion of *Pareto front layer* or *Pareto front rank*, which is the Pareto frontier of the data points once we exclude the Pareto-dominating points. For example, if we exclude the Pareto frontier shown in the step function in Figure 7, the Pareto frontier for the remaining points is the second Pareto front layer. We can exclude points in each iteration to obtain all the Pareto front layers. Figure 7 shows each Pareto front layer in a different color. At a high level, a greedy algorithm can use Pareto front layers to determine which points must be assigned to either treatment or control. The better the Pareto front rank, the greater the overall benefit of assigning those points to the treatment. We formalize this intuition in the following proposition that helps us link the CATE estimates to the policy performance under each objective:

Proposition 1. *Suppose that we have a policy π with M points assigned to the treatment and $N-M$ points assigned to the control. Consider a set of policies Π_c whose assignment is identical to the policy π, except for only one of the $N-M$ points that are assigned to the control condition. Suppose that the set of observations assigned to the control condition is $\mathcal{X}_c = \{X_1, X_2, \ldots, X_{N-M}\}$. For any $X_i \in \mathcal{X}_c$, let $\pi^{(X_i)}$ denote the policy that switches the treatment assignment for observation X_i, i.e., $\pi^{(X_i)}(x) = \mathbb{1}(\pi(x) = 1 \vee x = X_i)$. The Pareto frontier of $N-M$ policies in Π_c in terms of all K objectives are the policies that include the Pareto frontier of points $\{(\tau_{Y_1}(X_i), \tau_{Y_2}(X_i), \ldots, \tau_{Y_K}(X_i))\}_{i=1}^{N-M}$, where $\tau_{Y_j}(x)$ is the CATE for outcome Y_j when using the treatment compared to the control.*

Proof. Please see Web Appendix C.1. □

This proposition inspires us to propose a greedy algorithm called Greedy Front Elimination (GFE). The high-level intuition behind this algorithm is to start with the uniform control policy and switch one front layer at a time. Once the algorithm goes through all the front layers, the policy will completely switch to the uniform treatment policy. Figure 8 shows how this algorithm works at a given step. In this example, we have the Pareto front layer 25, and we choose all the points on this Pareto front layer and above (i.e., rank lower than or equal to 25) to receive the Skippable/Long ad format and the rest to receive the Non-Skippable/Short ad format. In line with Proposition 1, we choose the best candidates at every step to assign to the Skippable/Long ad format.

Algorithm 1 presents a formal description of the Greedy Front Elimination algorithm. The algorithm takes CATE estimates as inputs and returns a set of policies as the output. As shown in the while loop, we switch the policy for a Pareto front layer at any iteration. The fifth line in the algorithm presents a general equation for the policy that is defined based on a Pareto front layer. That is, the value of x does not need to be from our data, and we can determine the policy for any value of x. Specifically, if there is any point on a specific Pareto front layer (logical disjunction in line 5) that dominates CATE estimates for x (logical conjunction in line 5), the policy for x will

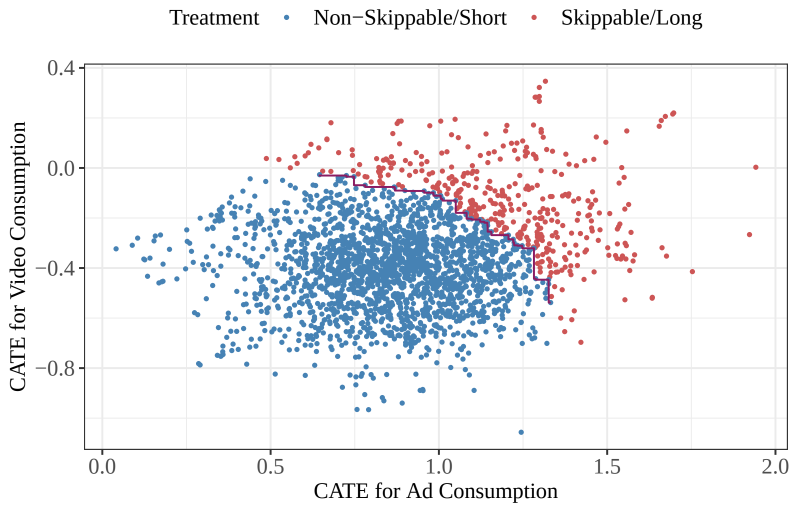


Figure 8. An illustration of the assignment policy in Greedy Front Elimination algorithm at Pareto front layer 25.

Algorithm 1 Greedy Front Elimination (GFE)

Input: $\{(\tau_{Y_1}(X_i), \tau_{Y_2}(X_i), \ldots, \tau_{Y_K}(X_i))\}_{i=1}^{N}$ ▷ CATE estimates for all outcomes

Output: Π^{GFE}

$j \leftarrow 1$
while $\mathcal{D} \neq \varnothing$ **do**
  $\mathcal{L}_j \leftarrow \mathrm{ParetoFront}(\mathcal{D})$ ▷ Indices for the Pareto front of $\mathcal{D}$
  $\mathcal{D} \leftarrow \mathcal{D} \setminus \mathcal{L}_j$
  $\pi_j^{\mathrm{GFE}}(x) \leftarrow 1 - \mathbb{1}\left(\bigvee_{i \in \mathcal{L}_j} \bigwedge_{k=1}^{K} \tau_{Y_k}(x) \leq \tau_{Y_k}(X_i)\right)$
  $j \leftarrow j + 1$
end while
$\Pi^{\mathrm{GFE}} \leftarrow \bigcup_{i=1}^{j} \pi_j^{\mathrm{GFE}}(x)$

---

be the control condition. Therefore, we can learn the set of policies on training data and apply it to separate test data.

5.2.2 Parameter-Agnostic Scalarization

The most common solution to the problem of multi-objective personalization is mapping multiple objectives into a single objective by weighting. This approach is known as "scalarization" in the literature on multi-objective optimization (Miettinen and Mäkelä 2002, Marler and Arora 2004). From an economic point of view, the weights can reflect the utility that the platform receives from each objective. The following proposition helps determine the policy for any given scalarization:

Proposition 2. *Suppose that the platform has K different objectives, each denoted by $\rho_{Y_j}(\pi)$ and*

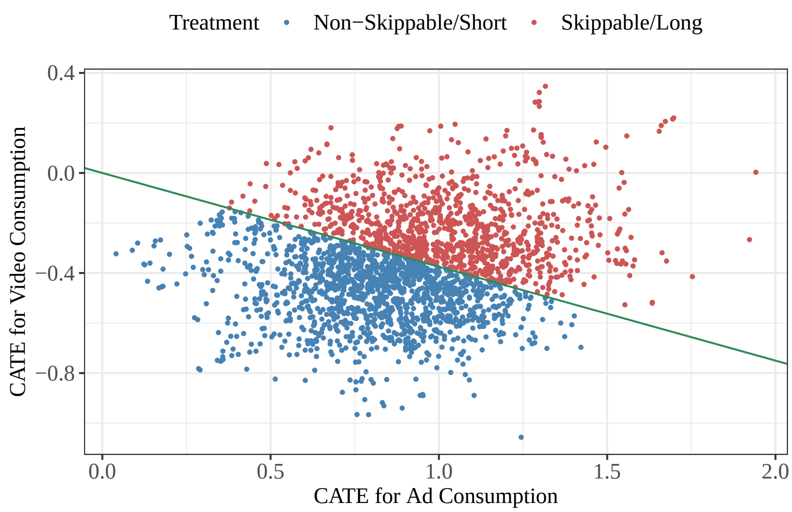


Figure 9. An illustration of the assignment policy for a set of given weights.

the platform wants to find the policy that optimizes the following joint objective:

$$\underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \rho_{Y_j}(\pi), \tag{5}$$

where β_j is the utility weight for the objective under outcome j. The optimal policy for each observation with the vector of covariates x will be the following:

$$\pi(x) = \mathbb{1}(\sum_{j=1}^{K} \beta_j \tau_{Y_j}(x) \geq 0), \tag{6}$$

where $\tau_{Y_j}(x)$ is the CATE for outcome Y_j when using the treatment compared to the control.

Proof. Please see Web Appendix C.2. □

Intuitively, the policies generated based on Proposition 2 split the space of CATE estimates by a hyperplane, where points on either side of the place are assigned to one of the policies. For example, in our problem with multiple objectives of Ad Consumption and Video Consumption, a line based on the weights in Equation (5) splits the CATE space into two parts. Figure 9 visualizes one of these policies on a random sample of our data points where the weights for ad and video consumption are 0.375 and 1, respectively. As shown in this figure, points above the line $\tau_v = -0.375\tau_a$ are assigned to the Skippable/Long ad format, and points below this line are assigned to the Non-Skippable/Short ad format.

The main drawback of scalarization as a multi-objective optimization approach is that the policy-maker needs to know the weights ex-ante. In many problems, the outcomes have vastly different units, making this a priori knowledge unattainable. In our context, the platform can use a variety of different weights and then trim the Pareto frontier of the full set. As such, the algorithm we use is parameter-agnostic. Let $\mathcal{B}$ denote the set of scalarization weights we want to consider. We can present the details of the algorithm as follows:

Algorithm 2 Parameter-Agnostic Scalarization

Input: $\{(\tau_{Y_1}(X_i), \tau_{Y_2}(X_i), \ldots, \tau_{Y_K}(X_i))\}_{i=1}^N$, $\mathcal{B}$
Output: Π^{PAS}

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

The output of Algorithm 2 is a set of policies that can each be evaluated by the platform. The platform can then use the performance of these policies to choose whichever best fits its objectives. Further, if the joint objective is non-linear, we can use a similar approach to split the space of CATE estimates. For example, in Figure 9, one could use a quadratic curve to arrive at the policy. However, we must notice that although the added complexity can add to the flexibility of the policy, it can create the well-known bias-variance trade-off.

5.3 Counterfactual Policy Evaluation

All the algorithms we have presented so far generate a set of policies. These policies have not been implemented in our data, but we need to evaluate what would have happened had the platform implemented these policies. As such, the question of evaluating a certain policy π becomes one of counterfactual policy evaluation. Since our CATE estimates are structural parameters, we can relatively compare the performance of a policy π with any given baseline policy. For example, let π_D denote the policy implemented in the data, i.e., $\pi_D(X_i) = W_i$. We can write:

$$\rho_Y(\pi) - \rho_Y(\pi_D) = \sum_{i=1}^N (\pi(X_i) - \pi_D(X_i))\, \tau_Y(X_i), \tag{7}$$

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

To address this challenge, we use an Inverse Propensity Scoring (IPS) estimator, which was first proposed by Horvitz and Thompson (1952) and defined as follows:

Definition 4. *IPS estimator can evaluate any outcome Y under the policy π in data $\mathcal{D}$ as follows:*

$$\hat{\rho}_Y^{\text{IPS}}(\pi; \mathcal{D}) = \frac{1}{N_{\mathcal{D}}} \sum_{i \in \mathcal{D}} \frac{\mathbb{1}\left(W_i = \pi(X_i)\right) Y_i}{e(W_i; \mathcal{D})}, \tag{8}$$

where $e(W_i; \mathcal{D})$ is the propensity score for treatment value W_i in observation i.

As shown in Equation (8), the IPS estimator is model-free 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 would have happened had the policy π been implemented.[9]

Further, because the IPS estimator is defined on the data $\mathcal{D}$, we can easily evaluate both the in-sample and out-of-sample performance of different policies. In particular, we randomly split our data into two sets, where 60% of the observations construct the training data $\mathcal{D}_{\text{Train}}$, and the remaining 40% constitute the test data $\mathcal{D}_{\text{Test}}$. We address the model-based error by performing CATE estimation and policy identification on the training data and evaluating its performance on separate held-out test data. Besides its robustness to model-based errors, our approach is useful as it mimics the practice of real-time policy-making, where the platform uses a batch of data to identify the policies and assign policies in real-time (test data). Thus, platforms can readily apply our framework.

5.4 Results

5.4.1 Raw Policy Comparison

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

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

policies. We present a short description of these policies as follows, and refer the reader to Web Appendix D for greater details:

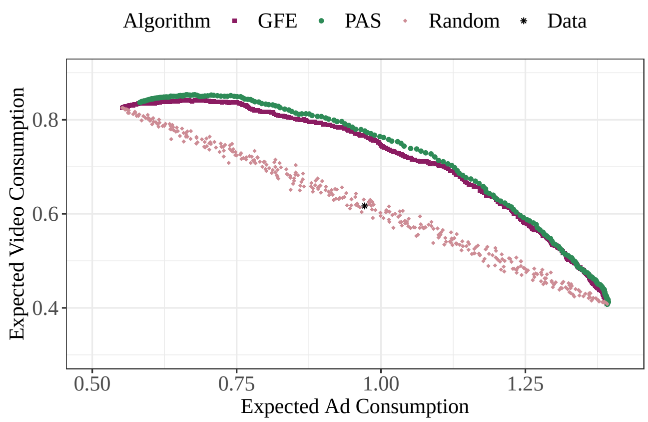
- *Greedy Front Elimination (GFE):* For all the points in the training data, we form the following set: $\mathcal{D}_{\text{Train}} = \{(\hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i))\}_i$. This set will be the input for Algorithm 1. The output of the algorithm will be a full set of policies that we denote by Π^{GFE}.
- *Parameter-Agnostic Scalarization (PAS):* For a given set of weights β_a and β_v, let $\beta_A \rho_A(\pi) + \beta_V \rho_V(\pi)$ denote the joint objective function for the platform. We can characterize this line with a single parameter α such that the objective is $\alpha \rho_A(\pi) + (1-\alpha)\rho_V(\pi)$. To mirror the GFE policies, we use 387 values of $\alpha \in \{0, 1/386, 2/386, \ldots, 1\}$ and perform Algorithm 2. The resulting output is a set of policies that we denote by Π^{PAS}.
- *Random:* As a simple benchmark, we consider a set of random policies for different proportions of each treatment. That is, for any proportion γ, we randomly assign γ fraction of units to the Skippable/Long condition. We generate 387 such random policies where $\gamma \in \{0, 1/386, 2/386, \ldots, 1\}$ and denote the full set of random policies by Π^{RND}.

We now have three different sets of policies: Π^{GFE}, Π^{SCL}, and Π^{RND}. We evaluate all these policies on both train and test data using Equation (8). In addition, we also report the mean estimates of Ad Consumption and Video Consumption in our data for each sample. We present all these results in Figure 10. As expected, the mean estimates from the data fall on the line for the class of random policies because our data set is also generated by a random policy. The most apparent result from these figures is that both GFE and Scalarization algorithms push the Pareto frontier of the two objectives compared to the random policy. Both policy sets are able to improve one objective without hurting the other objective too much. In particular, we find that the platform can substantially increase the expected ad consumption while keeping the expected video consumption almost the same. We interpret the exact gains in each dimension in the next section.

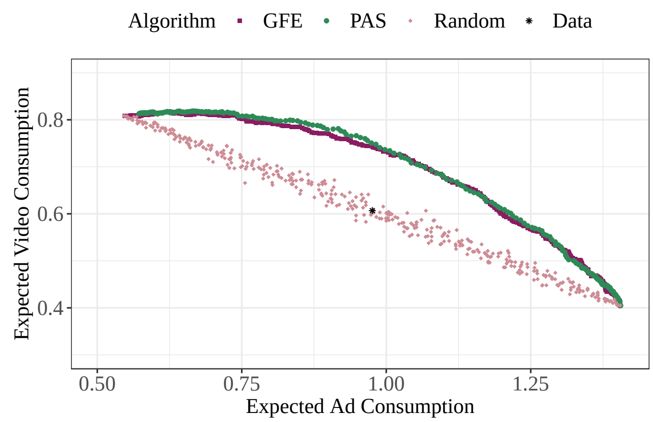
More specifically, we compare the performance of the two algorithms. The two algorithms perform very similarly, but the Parameter-Agnostic Scalarization algorithm performs slightly better, especially in generating policies that achieve higher expected video consumption (left side of figures). Finally, we note that both algorithms have some Pareto-dominated points in terms of the estimated outcomes. That is, there are two GFE policies where one achieves higher expected ad consumption and video consumption. This implies that we can trim these sets of policies and remove the Pareto-dominated policies. We do this in the next section.

5.4.2 Final Policy Comparison After Trimming

In this section, we trim both Π^{GFE} and Π^{PAS} so all the policies that are Pareto dominated by other policies in the same groups will be dropped. In particular, let $\mathcal{R}^{\text{GFE}}$ and $\mathcal{R}^{\text{PAS}}$ denote the estimated



(a) Training Data for Evaluation



(b) Test Data for Evaluation

Figure 10. Performance of the set of raw policies generated by different algorithms on the train and test data.

expected Ad Consumption and Video Consumption under policies of each group as follows:

$$\mathcal{R}^{\text{GFE}} = \left\{ \left(\hat{\rho}_A^{IPS}(\pi; \mathcal{D}_{\text{Train}}), \hat{\rho}_V^{IPS}(\pi; \mathcal{D}_{\text{Train}}) \right) \right\}_{\pi \in \Pi^{\text{GFE}}} \tag{9}$$

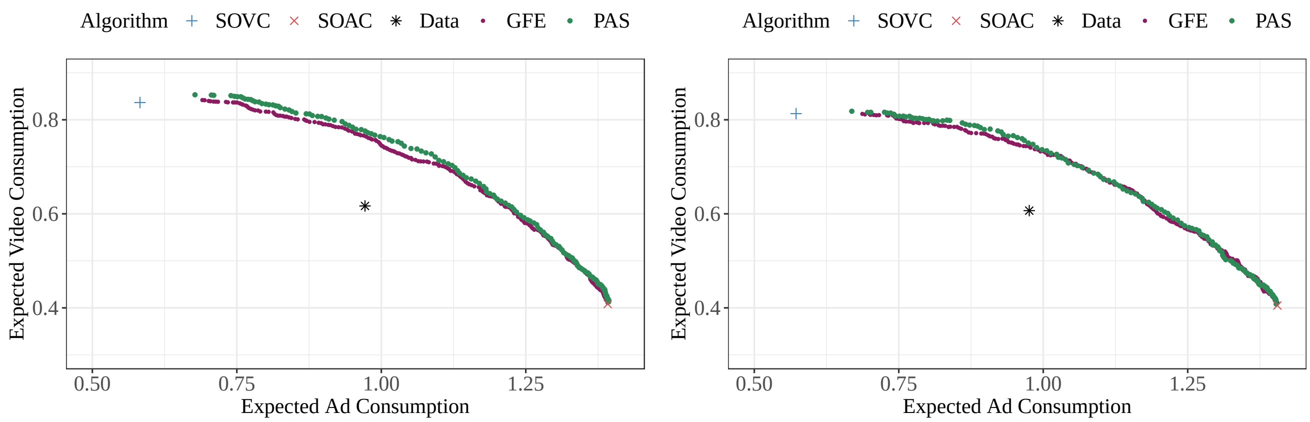
$$\mathcal{R}^{\text{PAS}} = \left\{ \left(\hat{\rho}_A^{IPS}(\pi; \mathcal{D}_{\text{Train}}), \hat{\rho}_V^{IPS}(\pi; \mathcal{D}_{\text{Train}}) \right) \right\}_{\pi \in \Pi^{\text{SCL}}} \tag{10}$$

We can trim the policies by removing the Pareto-dominated policies in $\mathcal{R}^{\text{GFE}}$ and $\mathcal{R}^{\text{PAS}}$. Let Π_f^{GFE} and Π_f^{PAS} denote the resulting sets of policies that only include the Pareto front of $\mathcal{R}^{\text{GFE}}$ and $\mathcal{R}^{\text{PAS}}$ respectively. Notice that we perform trimming on the training data to better mimic the real-world situation where platforms need to rely on batch data (train data) to identify the policy to deploy in real-time (test data). We want to compare these two sets of policies with the optimal single objective policies. To do so, let π^{SOAC} and π^{SOVC} denote the optimal personalized policies that only optimize Ad Consumption and Video Consumption, respectively. We can write:

$$\pi^{\text{SOAC}}(X_i) = \mathbb{1}(\hat{\tau}_A^{\text{Train}}(X_i) \geq 0) \tag{11}$$

$$\pi^{\text{SOVC}}(X_i) = \mathbb{1}(\hat{\tau}_V^{\text{Train}}(X_i) \geq 0) \tag{12}$$

Comparing multi-objective personalization algorithms (e.g., GFE and PAS) with single-objective personalization algorithms (e.g., SOAC and SOVC) allows us to quantify the value created by performing the multi-objective personalization task. We report the estimated outcomes under these groups of policies in Figure 11. We first note that the two single-objective personalized policies are the farthest away two points in these graphs in terms of expected ad and video consumption. As discussed earlier, the reason behind this pattern is the conflict between the treatment effects of two outcomes. In particular, observations with positive CATE estimates for ad consumption



(a) Training Data for Evaluation

(b) Test Data for Evaluation

Figure 11. Performance of the set of the final set of Pareto frontier policies generated by different algorithms on the train and test data.

generally have a negative CATE for the video consumption outcome.

The stark contrast between the performance of the two single-objective policies serves as the motivation for using multi-objective personalized policies whose main goal is improving one objective without hurting the other. However, it is important to notice that a platform cannot achieve all the points on the Pareto frontiers shown in Figure 11. This is because the platform can select only one policy. The value of multi-objective personalization is in providing a complete picture for a policy-maker to choose one of the policies on the Pareto frontier that best achieves their objectives. To that end, we find three notable results:

- *High Video Consumption, Medium Ad Consumption:* From Figure 11a, we see that the policy-maker can choose a variety of policies with great video consumption performance while improving on ad consumption. For example, the policy-maker can choose one of the Parameter-Agnostic Scalarization policies with $\alpha = 95/386$ that results in 4.8% lower video consumption compared to the Single-Objective Video Consumption (SOVC) policy, while increasing ad consumption by 60.1% on the training data. When evaluating the performance of this policy ($\pi^{\text{PAS}}_{95/386}$) with that of the single-objective video consumption (π^{SOVC}) on the test data, we find that it will result in a drop of 4.5% in video consumption while increasing ad consumption by 61.0% (from 0.57 to 0.92, or alternatively from 8.58 to 13.82 seconds).
- *High Ad Consumption, Medium Video Consumption:* On the right end of Figure 11a, the policy-maker can choose a policy from Scalarization with $\alpha = 140/386$ that achieves a 50.6% improvement in video consumption compared to the Single-Objective Ad Consumption (SOAC) policy, while only losing 11.9% in ad consumption. On the test data, the policy $\pi^{\text{PAS}}_{140/386}$ performs 47.1% better in terms of video consumption that the (π^{SOAC}) policy, at the expense of 13.3% worse performance in terms of ad consumption.

- *High Video Consumption, 15 Seconds Ad Consumption:* A useful feature of multi-objective personalization is that we can fix a value for one objective and examine the performance in terms of the other objective. Since ad consumption cannot technically be more than 15 seconds in the Non-Skippable/Short ad condition, setting ad consumption to 15 seconds would be a reasonable objective. We find that the GFE policy at Pareto front layer 145 achieves 15 seconds of expected ad consumption. We compare the performance of the GFE policy π_{145}^{GFE} with the two single-objective policies. Compared to the Single-Objective Video Consumption (SOVC) policy, it improves the expected ad consumption by 75.4%, while only reducing the video consumption by 10.2%, as measured on the test data. On the other end, policy π_{145}^{GFE} improves video consumption by 80.3% compared to the Single-Objective Ad Consumption (SOAC) policy, while losing 28.6% in ad consumption (equivalent of 6 seconds).

Together, we find that multi-objective personalization results in substantial gains in one objective without sacrificing too much in the other objective. In Web Appendix E, we consider alternative policy evaluation approaches and show the same qualitative results. Intuitively, multi-objective 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 and obtain the Pareto frontier, and then decide which policy on the Pareto frontier is more desirable.

Finally, we demonstrate that the platform can create substantial value by using multi-objective personalization, even in a setting with an almost perfect substitution between the two objectives. As shown earlier, for over 96% of the data points in our data, we observe some degree of substitution between Ad Consumption and Video Consumption. The gains can be significantly larger when the two objectives are less in conflict with each other. To demonstrate this point, we focus on another set of objectives in Web Appendix F: (1) Second 15 Complete, and (2) Video Consumption. Since many platforms charge advertisers once their ad is watched for 15 seconds (e.g., Facebook), using these two objectives for multi-objective personalization is reasonable for profit-maximizing platforms. We present the results of this practice in Web Appendix F and document substantial gains from a multi-objective personalization policy.

6 Implications

6.1 Implications for Video Advertising Platforms

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 ad and video consumption and showed that the platform could create value using our multi-objective personalization framework.

Although we focused on this set of objectives in our empirical context, our framework is fairly general, and the platform can apply it with a different set of objectives. For example, some platforms may be interested in increasing the rate at which the user reaches a certain point within the ad because they charge advertisers based on that rule. In our study, we can consider the 15-second threshold and perform multi-objective personalization for *Second 15 Complete* and *Video Consumption* as our main outcomes of interest (please see Web Appendix F 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 their objectives, our paper has important market design implications for video advertising platforms. These platforms generally sell ads through auctions. Any auction is characterized by an allocation rule and a payment rule. Our paper highlights why the allocation rule should not be only based on the ad performance but also on the externality it imposes on the system. Prior literature on advertising auctions has studied different forms of ad allocation that capture the externality an ad exposure imposes on other ads (Wilbur et al. 2013, Kar et al. 2015, 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.

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 the payment rule in these problems. That is, the platform needs to decide when to charge the advertisers. Some platforms use cutoff-based rules where the advertiser is charged for skippable ads if the user reaches the Second 30 of the ad. Part of the reason for having these rules in place is to account for the externalities an ad exposure can impose on content creators. Given the substitution between ad and video consumption, our findings suggest that a consumption-based payment rule can better account for these externalities. Furthermore, designing an auction with clearer guarantees under a consumption-based payment rule would be easier than in environments with arbitrary cutoff-based rules.

6.2 Implications for Advertisers and Content Creators

Although the main implications of our work are for platforms and market designers, our findings share important insights for advertisers and content creators. One of the decisions that advertisers have to make is to decide on the skippability and length of their ad. In many cases, like the ad in our experiment, they create more than one ad version, which raises the question of in which context each ad version performs better. We highlight that the impact of Skippable/Long ads relative to Non-Skippable/Short ads highly depends on the ad itself. In particular, we find that the presence of the skip option lowers immediate ad avoidance, which results in a higher 15-second completion rate under Skippable/Long ads compared to Non-Skippable/Short ads in our context. Hence, our results challenge the common understanding that skippability always reduces ad consumption and suggest that advertisers should run experiments to test the impact of ad formats on ad-related outcomes.

Our paper offers insights for content creators by establishing a strong substitution pattern between ad and video consumption. Although content creators know that the presence of ads reduces video consumption, the magnitude of this substitution has important implications for content creators who determine the types of ads that can compete for their slots. Further, our analysis of heterogeneity suggests different substitution patterns at different times of the day. Content creators can use this information to customize the ads at different times of the day.

7 Conclusions

Content-streaming platforms rely on digital video ads for monetization. These video ads largely vary in two features: length and skippability. The two common video ad formats are (1) Non-Skippable/Short ads that are relatively short in length (e.g., 15 seconds), but the user has to fully watch them to continue their session, and (2) Skippable/Long ads that are relatively long but users can skip them after a few seconds. In this paper, we partner with the video advertising platform `vdo.ai` and run a field experiment where we randomly assign users to Non-Skippable/Short and Skippable/Long versions of an ad for an identical product. We document a substitution pattern between ad and video consumption using experimental data. Although the Skippable/Long ad format substantially increases ad consumption compared to the Non-Skippable/Short ad format, it decreases the consumption of the subsequent video, on average. This substitution pattern between the two outcomes creates a challenge for a platform that wants to increase both ad and video consumption. We estimate the conditional average treatment effects (CATE) and find that the substitution pattern exists even at the individual user level.

Motivated by this challenge, we develop algorithms for multi-objective personalization whose goal is to develop a set of policies on the Pareto frontier of expected ad and video consumption outcomes. These algorithms exploit the magnitude of the substitution at the individual level to assign individuals to policies. We find that multi-objective personalized policies improve the outcome in one dimension compared to single-objective personalized policies without hurting the outcome in the other dimension. In particular, we show that compared to a single-objective personalized policy that only optimizes video consumption, there is a policy on the identified Pareto frontier that improves ad consumption by 61% while only reducing video consumption by 4%. Likewise, we document that compared to the single-objective personalized policy that only optimizes ad consumption, there is a multi-objective personalized policy that increases video consumption by 47% while only decreasing the ad consumption outcome by 13%. We discuss the implications and how the platform can use our framework for optimal decision-making in real-time.

Our research offers several contributions to the literature. From a substantive standpoint, we examine the relationship between two crucial metrics in the industry: ad consumption and video consumption. Using a field experiment, we determine the causal relationship between the two and quantify the substitution effect. Although context-specific, our findings challenge two common beliefs about video ads: (1) all users receive a negative utility of consumption, and (2) skippability always results in a lower completion rate of the same length of the ad. Methodologically, we propose a framework for multi-objective personalization and introduce two classes of algorithms, non-parametric and parametric. Our approach draws upon the principles of multi-objective optimization to learn personalized policies that optimize multiple objectives. A key insight of our paper is that even when single-objective personalization does not generate any gains relative to a uniform policy, multi-objective personalization can exploit cross-outcome effects and develop policies that generate substantial gains in multiple outcomes. From a practical standpoint, our framework provides flexibility for policymakers and managers in balancing multiple outcomes and selecting policies that align with their goals and can be broadly applied to a variety of applied problems.

Nevertheless, our paper has certain limitations that serve as excellent avenues for future research. First, we only focus on Skippable/Long and Non-Skippable/Short versions of only one ad. Given that ad consumption highly depends on the ad itself, an important avenue for future research would be to quantify the determinants of treatment effect heterogeneity across ads. Second, a limitation of our data is the lack of differentiation between the videos. Future studies can incorporate rich and high-dimensional video information and document the heterogeneity in the main effects across videos. 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, although our research offers implications for the design of video advertising auctions, exploring the theoretical properties of different types of auctions is beyond the scope of our research. Future work can theoretically study the auction design problem in video advertising auctions and examine its differences from other types of advertising auctions.

Disclosure Statement

The author certifies that he has no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The author has no funding to report.

References

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

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

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

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360.

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

Biswas, S. (2020). *Investigating the effects of including discount information in advertising*. PhD thesis, The University of Chicago.

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

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

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

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

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

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

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

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

IAB (2022). Digital Video Ad Spend Increased 49% in 2021 and Expected to Reach Nearly $50 Billion in 2022, According to IAB's 2021 Video Ad Spend and 2022 Outlook Report .

Jeon, Y. A., Son, H., Chung, A. D., and Drumwright, M. E. (2019). Temporal certainty and skippable in-stream commercials: Effects of ad length, timer, and skip-ad button on irritation and skipping behavior. *Journal of Interactive Marketing*, 47(1):144–158.

Kar, W., Swaminathan, V., and Albuquerque, P. (2015). Selection and ordering of linear online video ads. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 203–210.

Kohavi, R. and Thomke, S. (2017). The surprising power of online experiments. *Harvard business review*, 95(5):74–82.

Lambrecht, A. and Tucker, C. (2013). When does retargeting work? information specificity in online advertising. *Journal of Marketing research*, 50(5):561–576.

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

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

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

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

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

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

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

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

Rafieian, O. (2022). Optimizing user engagement through adaptive ad sequencing. Technical report, Working paper.

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

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

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

Sahni, N. S. and Nair, H. S. (2020). Sponsorship disclosure and consumer deception: Experimental evidence from native advertising in mobile search. *Marketing Science*, 39(1):5–32.

Schumann, D. W., Petty, R. E., and Scott Clemons, D. (1990). Predicting the effectiveness of different strategies of advertising variation: A test of the repetition-variation hypotheses. *Journal of Consumer Research*, 17(2):192–202.

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

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

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

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

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

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

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

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

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

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

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

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

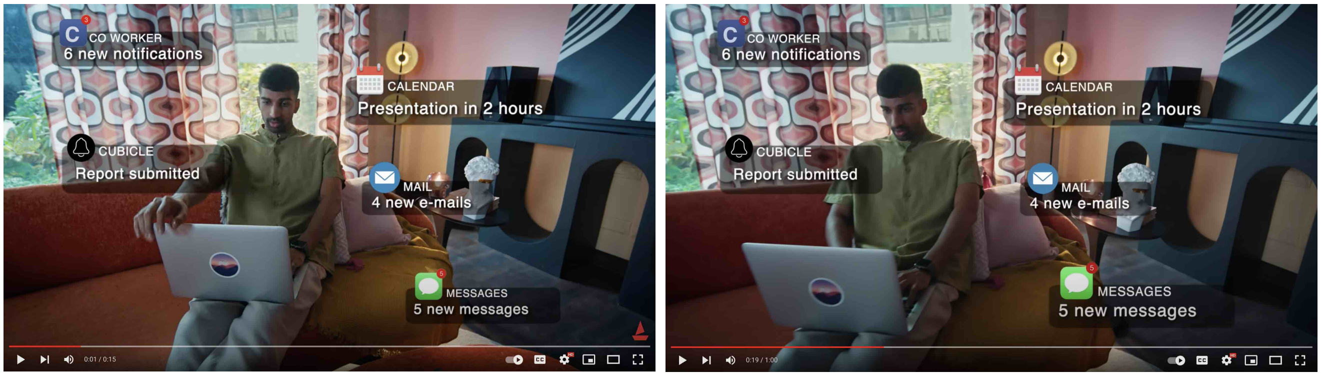
Wilbur, K. C., Xu, L., and Kempe, D. (2013). Correcting audience externalities in television advertising. *Marketing Science*, 32(6):892–912.

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

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

Web Appendix

A Screenshots of `boAt` Ads



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

Figure A1. Screenshots of Non-Skippable/Short and Skippable/Long ads of boAt product used in the experiment.

B Randomization Check

In this section, we use the pre-treatment variables to check whether randomization in our experiment has been implemented properly. As discussed earlier, we use a 45%-45%-10% split for our three treatment arms, such that both Skippable/Long and Non-Skippable/Short formats receive an equal 45% share, and the No-Ad condition receives 10% of the total traffic. In our data, we see 25,021 observations in the Skippable/Long ad condition (44.16%), 26,402 observations in the Non-Skippable/Short ad condition (46.60%), and 5,239 observations in the No-Ad condition (9.25%). These numbers are consistent with the splitting goal of the experiment.

We now verify whether the treatment is properly randomized. To do so, we need to check if there is any discrepancy in the distribution of the pre-treatment covariates across treatment conditions. Because our main goal is to compare the two ad formats, we mainly focus on two treatment arms: Skippable/Long and Non-Skippable/Short ad formats. For each pre-treatment variable X, let $\mu_{X,0}$ and $\mu_{X,1}$ denote the population means of that variable in the Non-Skippable/Short and Skippable/Long conditions, respectively. If randomization has been done properly, we will fail to reject the following null hypothesis: $H_0 : \mu_{X,0} = \mu_{X,1}$.

In our setting, all the pre-treatment variables are categorical. As a result, each pre-treatment variable X is a dummy for a specific subcategory. We conduct three separate tests for randomization checks. First, we conduct Fisher's exact test for each subcategory. Since we run multiple hypotheses, we expect a fraction of them to be significant even if the null hypothesis is true. Of

837 separate tests conducted, only 8 rejected the null hypothesis. After adjusting for multiple hypothesis testing using the Benjamini-Hochberg approach (Benjamini and Hochberg 1995), no adjusted p-value was below 0.05.[10]

Second, we use the measure of *Standardized Bias (SB)*, which is commonly used in the literature to assess covariate balance. Standardized Bias is equal to the absolute difference between the means of two groups divided by the standard deviation of the covariate for the pooled sample. The common norm in the literature is to consider a Standardized Bias below 0.2 or 0.1 as evidence for covariate balance (McCaffrey et al. 2013). In our setting, we find that the maximum Standardized Bias was 0.026, which indicates that we have a covariate balance for all the pre-treatment covariates using this approach.

Finally, we use a regression approach to regress the treatment assignment on all the pre-treatment variables. If randomization has been done properly, the pre-treatment variables will have no predictive power in explaining the treatment assignment. We can statistically test that by using the F-test of the regression model. We find that the F-statistic is equal to 1.02 with a p-value of 0.32, which indicates that the pre-treatment variables have no predictive power in predicting the treatment assignment and provides evidence for the validity of randomization in our study.

C Proofs

C.1 Proof for Proposition 1

Proof. Let $\mathcal{X}_{c,p}$ denote the Pareto frontier of points $\{(\tau_{Y_1}(X_i), \tau_{Y_2}(X_i), \dots, \tau_{Y_K}(X_i))\}_{i=1}^{N-M}$. For the proof, it suffices to show two things: (1) any policy $\pi^{(x)}$ such that $x \in \mathcal{X}_c \setminus \mathcal{X}_{c,p}$ is Pareto dominated, and (2) any policy $\pi^{(x)}$ such that $x \in \mathcal{X}_{c,p}$ is on the Pareto frontier. We prove the first one by contradiction. Suppose that policy $\pi^{(x)}$ such that $x \in \mathcal{X}_c \setminus \mathcal{X}_{c,p}$ is on the Pareto front. For each outcome Y_j, we can write the following about the performance of this policy $\pi^{(x)}$:

$$\rho_{Y_j}\left(\pi^{(x)}\right) = \rho_{Y_j}(\pi) + \tau_{Y_j}(x) \tag{13}$$

If $x \in \mathcal{X}_c \setminus \mathcal{X}_{c,p}$, we know that there exists a point $x' \in \mathcal{X}_{c,p}$ for which we have $\tau_{Y_j}(x') \geq \tau_{Y_j}(x)$ for all j. This implies that there is a policy $\pi^{(x')}$ that dominates $\pi^{(x)}$ in terms of the performance for each objective j, that is, $\rho_{Y_j}(\pi^{(x')}) \geq \rho_{Y_j}(\pi^{(x)})$ for all j. Therefore, policy $\pi^{(x)}$ such that $x \in \mathcal{X}_c \setminus \mathcal{X}_{c,p}$ is not on the Pareto frontier. To show the second part, we use Equation (13), which indicates that if $x \in \mathcal{X}_{c,p}$ is not Pareto dominated in terms of CATE, $\pi^{(x)}$ is not Pareto dominated by any other policy, as there is a one-to-one mapping between the $\mathcal{X}_c$ and Π_c. $\square$

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

C.2 Proof for Proposition 2

Proof. For any vector of covariates x, the joint objective in Equation (5) can be written as follows:

$$\underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \left(\rho_{Y_j}(0) + \tau_{Y_j}(x)\pi(x)\right). \tag{14}$$

Since the term $\rho_{Y_j}(0)$ is policy-invariant, the optimal policy will be the one that maximizes $\sum_{j=1}^{K} \beta_j \tau_{Y_j}(x)\pi(x)$, which is equal to the treatment if $\sum_{j=1}^{K} \beta_j \tau_{Y_j}(x) \geq 0$. □

D Details of Policies Defined in §5.4.1

In this section, we present a detailed and more formal version of the policies defined in §5.4.1. We use our CATE estimates of Ad Consumption and Video Consumption on the training data to identify the following sets of policies:

- *Greedy Front Elimination (GFE):* For all the points in the training data, we form the following set: $\mathcal{D}_{\text{Train}} = \{(\hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i))\}_i$. This set will be the input for Algorithm 1. For all the points in $\mathcal{T}$, we find all the Pareto front layers. Using our estimates, we find a total of 387 Pareto front layers in our training data. For each particular Pareto front layer j, let $\mathcal{L}_j$ denote the set of indices for points in that Pareto front layer. We denote the GFE policy corresponding to Pareto front layer j by π_j^{GFE} and define it as follows:

$$\pi_j^{\text{GFE}}(X_i) = 1 - \mathbb{1}\left(\bigvee_{k \in L_j} \left(\hat{\tau}_A^{\text{Train}}(X_i) \leq \hat{\tau}_A^{\text{Train}}(X_k) \wedge \hat{\tau}_V^{\text{Train}}(X_i) \leq \hat{\tau}_V^{\text{Train}}(X_k)\right)\right), \tag{15}$$

  where the right-hand side only takes value one if the CATE estimates for X_i are not Pareto dominated by the Pareto front layer j. We further define the full set of policies as $\Pi^{\text{GFE}} = \bigcup_j \pi_j^{\text{GFE}}$, where the union is defined over all 387 Pareto front layers.

- *Parameter-Agnostic Scalarization (PAS):* For a given set of weights β_a and β_v, let $\beta_A \rho_A(\pi) + \beta_V \rho_V(\pi)$ denote the joint objective function for the platform. The resulting policy will be determined by the line $\tau_V = -(\beta_A/\beta_V)\tau_A$. As such, we can represent each linear policy with a single parameter α such that $0 \leq \alpha \leq 1$ and $\tau_V = -(\alpha/(1-\alpha))\tau_A$. Since the platform wants higher ad and video consumption, it is reasonable to search only among positive values of α. To mirror the GFE policies, we use 387 values of $\alpha \in \{0, 1/386, 2/386, \ldots, 1\}$. For each α, we define the corresponding scalarization policy as π_α^{SCL} as follows:

$$\pi_\alpha^{PAS}(X_i) = \mathbb{1}\left(\hat{\tau}_V^{\text{Train}}(X_i) \geq \frac{\alpha}{1-\alpha}\hat{\tau}_A^{\text{Train}}(X_i)\right) \tag{16}$$

We define the full set of Scalarization policies as Π^{PAS}.

- *Random:* As a simple benchmark, we consider a set of random policies for different proportions of each treatment. Let π_γ^{RND} denote the random policy where the proportion of Skippable/Long ad conditions is γ. We can write:

$$\pi_\gamma^{RND}(X_i) = \mathbb{1}\left(u_i \leq \gamma\right), \tag{17}$$

where $u_i \sim U[0,1]$. We generate 387 such random policies where $\gamma \in \{0, 1/386, 2/386, \dots, 1\}$. We denote the full set of random policies by Π^{RND}.

E IPS Estimator with Estimated Propensity Scores

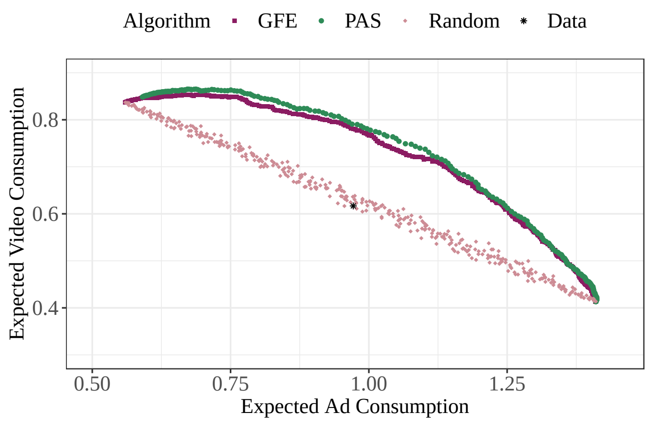
In §5.3, the estimator in Definition 4 uses known weights to evaluate to performance of policies given data. We used this approach because randomization happened by design, and the probability of users receiving Skippable/Long vs. Non-Skippable/Short is the same. Since we focus only on the sample of data where either one of these two treatment conditions has been implemented and drop the No-Ad condition, the known propensity would be 0.5, which gives us an inverse propensity score of 2. In this section, we estimate the propensity scores using the pre-treatment variables, so if there is any selection on observables, it would be captured.

To perform this task, we use a predictive XGBoost model to estimate the probability that a user receives the Skippable/Long ad version, given their pre-treatment observables, including time, location, and the operating system. We estimate these propensity scores that feed them to the denominator of Equation (8). Figure A2 presents the performance of both raw and final (after trimming) policies when using IPS with estimated propensity scores as the policy evaluation strategy. The results show the same pattern as those presented in the main text of the paper. Thus, all the insights remain the same once we estimate propensity scores.

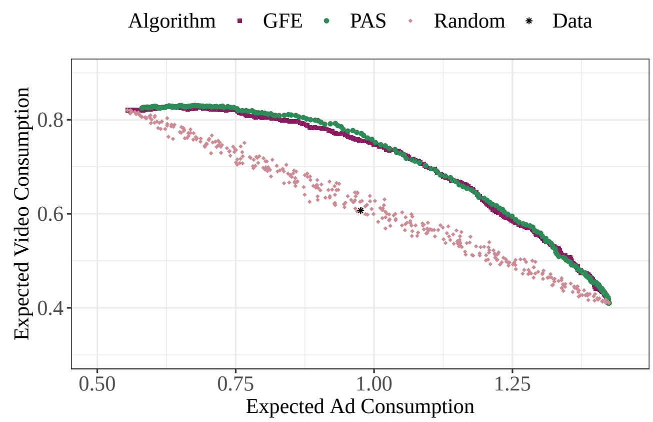
F Multi-Objective Personalization with Different Outcomes

In this section, we perform the multi-objective personalization framework when optimizing *Second 15 Complete* and *Video Consumption*. This approach reflects the joint utility of many video advertising platforms that charge advertisers when the user reaches a certain point within the ad. That is, if the user reaches a certain point in the ad, the advertiser has to pay even if the user later skips the ad. The cutoff rule varies across platforms ranging from 15 to 30 seconds. Thus, a natural problem objective for platforms is to maximize the ad revenue by having more people reach the cutoff point while keeping video consumption high.

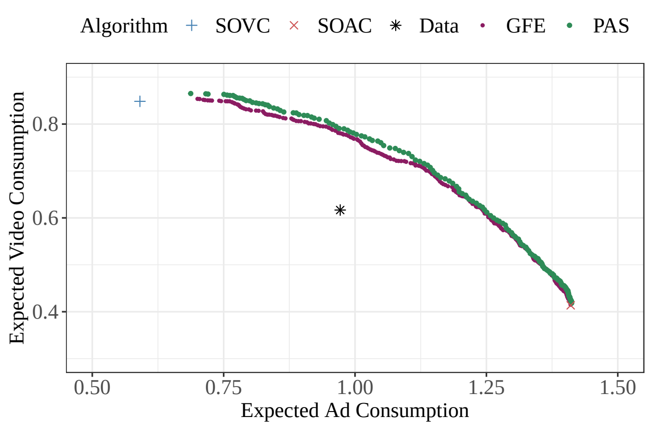
In §3, we presented the average treatment effect on Second 15 Complete as the outcome variable. We showed that users in the Skippable/Long ad condition are more likely to reach the 15^{th} second



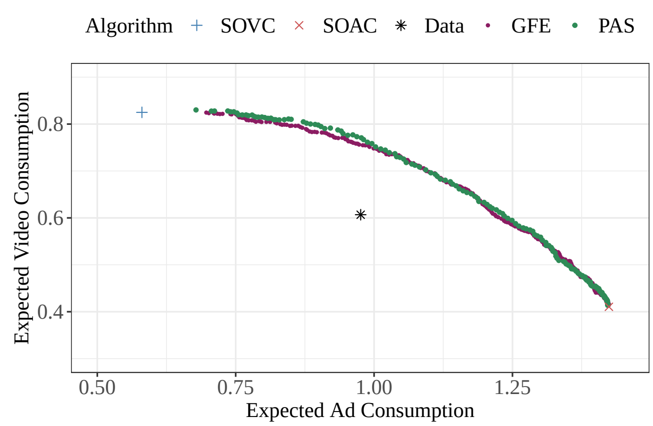
(a) Training Data for Raw Evaluation



(b) Test Data for Raw Evaluation



(c) Training Data for Evaluation After Trimming



(d) Test Data for Evaluation After Trimming

Figure A2. Performance of the set of raw and final policies generated by different algorithms on the train and test data using the IPS estimator as the counterfactual policy evaluation algorithm when using estimated propensity scores.

of the ad. However, the magnitude of the treatment effect is smaller compared to the effects on Ad Consumption. In particular, we do not expect a natural substitution pattern between Second 15 Complete and Video Consumption. Therefore, we expect that applying multi-objective personalization to this problem creates more substantial value.

We first estimate the CATE on Second 15 Complete using Causal Forests, with a 10-fold cross-validation. We plot the CATE on Video Consumption against CATE on Second 15 Complete for a random sample of observations in our and present the results in Figure A3 to see the extent to which the two outcomes are in conflict. Unlike the case with Ad Consumption and Video Consumption, we note that the sign of the CATE estimates is the same for a large portion of observations. Over 41% of all observations have the same signs of CATE estimates, which means that the treatment assignment for these observations is clear: units with positive CATE estimates on both outcomes receive the Skippable/Long ad, whereas units with negative CATE estimates

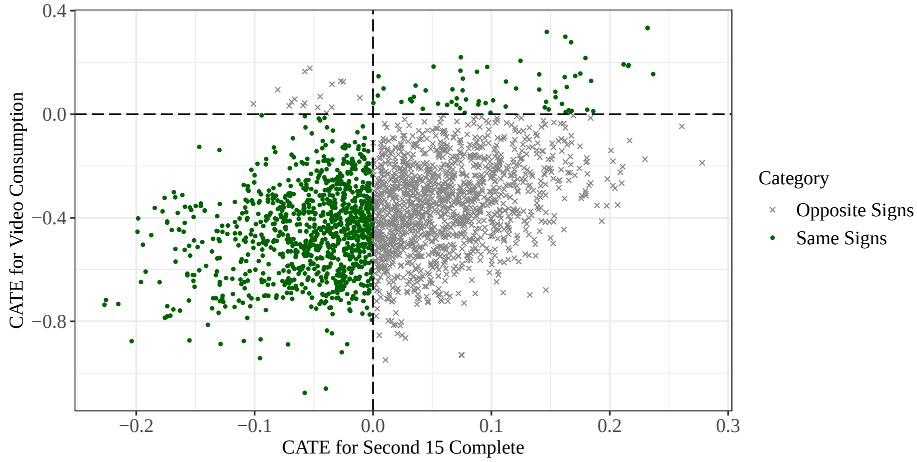
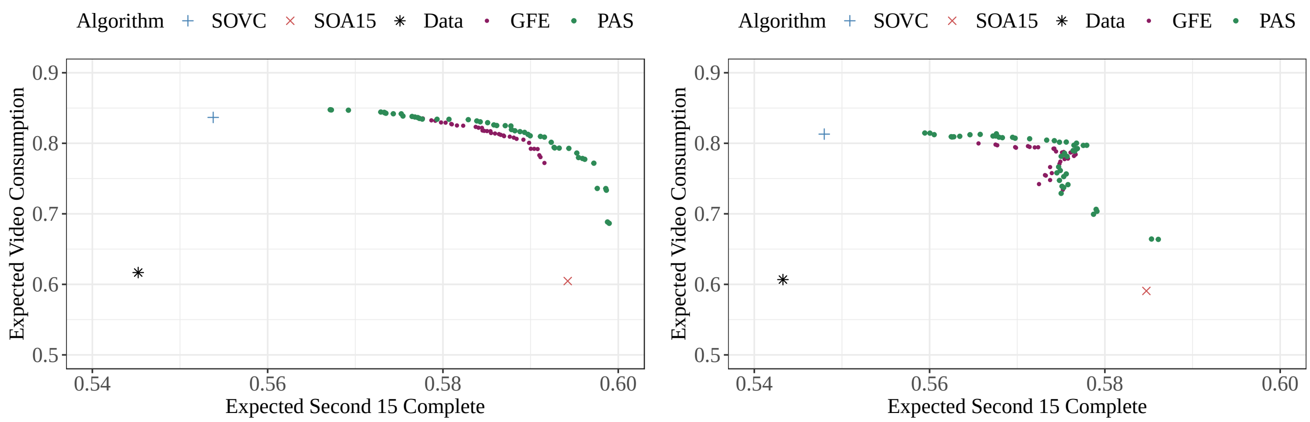


Figure A3. Scatter plot of CATE estimates for Video Consumption (in quarters) against CATE estimates for Second 15 Complete.

on both outcomes receive Non-Skippable/Short ad. Further, we find a positive correlation of 0.38 between CATE on Second 15 Complete and CATE on Video Consumption. This confirms our initial intuition that multi-objective personalization would be valuable in this setting.

We then apply both Greedy Front Elimination and Parameter-Agnostic Scalarization algorithms to this set of CATE estimates to generate a set of policies under each algorithm. We evaluate the performance of these policies using the IPS estimator presented in Definition 4 on the training data and trim the ones that are Pareto dominated. We consider three benchmark policies: (1) a Single-Objective Second 15 Complete policy that we abbreviate with SOA15, (2) a Single-Objective Video Consumption policy (SOVC as before), and (3) the policy implemented in our data.

We present the performance of all these policies using the IPS estimator when evaluated on the training and test data in Figure A4. As shown in this figure, the multi-objective personalization framework can offer a solution that achieves single-objective optimal outcomes in one dimension almost at no loss in the other dimension. More specifically, there is a PAS policy with $\alpha = 319/422$, which achieves great outcomes. Compared to the Single-Objective Second 15 Complete (SOA15) policy, this multi-objective personalized policy improves Video Consumption by 36.6%, while reducing the Second 15 Complete rate by 1.4% on the training data. On the test data, the gain in Video Consumption is 35.5%, at the same loss of 1.4% of the Second 15 Complete rate. The same policy compares well with the Single-Objective Video Consumption (SOVC) policy. On the training data, this multi-objective personalized policy generated by the PAS algorithm with $\alpha = 319/422$ increases the Second 15 Complete rate by 5.8% relative to the SOVC policy, while reducing Video Consumption by 1.3%. On the test data, the gain in the Second 15 Complete rate is 5.3%, and the drop in Video Consumption is 1.6%. Together, our results show that the



(a) Training Data for Evaluation

(b) Test Data for Evaluation

Figure A4. Performance of the set of the final set of Pareto frontier policies generated by different algorithms on the train and test data.

multi-objective personalization framework can be applied to a variety of settings and generate gains beyond the single-objective personalization framework.

References

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

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