Web Appendix for Multi-Objective Personalization of Marketing Interventions

Omid Rafieian — Cornell University
Anuj Kapoor — University of Missouri
Amitt Sharma — VDO.AI

August 6, 2024

Web Appendix

A Supplementary Materials for Multi-Objective Personalization Framework

A.1 Proof for Proposition 1

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

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

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

$$1 - \pi^*(x) = \lambda\big(1 - \pi_1(x)\big) + (1-\lambda)\big(1 - \pi_2(x)\big) \tag{2}$$

Now, for any outcome Y_i, we can write:

$$
\begin{aligned}
\rho(\pi^*) =& \mathbb{E}\left[\pi^*(X_i)Y_i(1) + \big(1-\pi^*(X_i)\big)Y_i(0)\right] \\
=& \mathbb{E}\left[\big(\lambda\pi_1(X_i) + (1-\lambda)\pi_2(X_i)\big)Y_i(1) + \Big(\lambda\big(1-\pi_1(X_i)\big) + (1-\lambda)\big(1-\pi_2(X_i)\big)\Big)Y_i(0)\right] \\
=& \mathbb{E}\left[\lambda\Big(\pi_1(X_i)Y_i(1) + \big(1-\pi_1(X_i)\big)Y_i(0)\Big)\right] + \\
& \mathbb{E}\left[\Big((1-\lambda)\Big(\pi_2(X_i)Y_i(1) + \big(1-\pi_2(X_i)\big)Y_i(0)\Big)\Big)\right] \\
=& \lambda\rho(\pi_1) + (1-\lambda)\rho(\pi_2),
\end{aligned}
\tag{3}
$$

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

A.2 Proof for Proposition 2

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

$$
\begin{aligned}
\pi^S_\beta &= \underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_k \rho_j(\pi) \\
&= \underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \mathbb{E}\left[\pi(X_i)Y_i^{(j)}(1) + (1-\pi(X_i))Y_i^{(j)}(0)\right] \\
&= \underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \mathbb{E}\left[Y_i^{(j)}(0) + \pi(X_i)\left(Y_i^{(j)}(1) - Y_i^{(j)}(0)\right)\right] \\
&= \underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \left(\rho_j(0) + \tau_j(x)\pi(x)\right) \\
&= \underset{\pi}{\operatorname{argmax}} \sum_{j=1}^{K} \beta_j \tau_j(x)\pi(x),
\end{aligned}
\tag{4}
$$

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

A.3 Proof for Proposition 3

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

$$
\begin{aligned}
\sum_{j=1}^{K}\beta_j\hat{\rho}_j^{IPS}(\pi;\mathcal{D}) &= \sum_{j=1}^{K}\beta_j\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}\left(\frac{\pi(X_i)W_i}{e(X_i)}+\frac{(1-\pi(X_i))(1-W_i)}{1-e(X_i)}\right)Y_i^{(j)} \\
&=\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}\left(\frac{\pi(X_i)W_i}{e(X_i)}+\frac{(1-\pi(X_i))(1-W_i)}{1-e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right)
\end{aligned}
\tag{5}
$$

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

$$
\begin{aligned}
\operatorname*{argmax}_{\pi}\sum_{j=1}^{K}\beta_j\hat{\rho}_j^{IPS}(\pi) &= \operatorname*{argmax}_{\pi}\sum_{j=1}^{K}2\beta_j\hat{\rho}_j^{IPS}(\pi)-\sum_{j=1}^{K}\beta_j\left(\hat{\rho}_j^{IPS}(1)+\hat{\rho}_j^{IPS}(0)\right) \\
&= \operatorname*{argmax}_{\pi}\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}2\left(\frac{\pi(X_i)W_i}{e(X_i)}+\frac{(1-\pi(X_i))(1-W_i)}{1-e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right) \\
&\qquad -\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}\left(\frac{W_i}{e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right) \\
&\qquad -\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}\left(\frac{(1-W_i)}{1-e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right) \\
&= \operatorname*{argmax}_{\pi}\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}\left(\frac{(2\pi(X_i)-1)W_i}{e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right) \\
&\qquad +\frac{1}{N_{\mathcal{D}}}\sum_{i\in\mathcal{D}}\left(\frac{(1-2\pi(X_i))(1-W_i)}{1-e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right) \\
&= \operatorname*{argmax}_{\pi}\,(2\pi(X_i)-1)\left(\frac{W_i}{e(X_i)}-\frac{(1-W_i)}{1-e(X_i)}\right)\left(\sum_{j=1}^{K}\beta_jY_i^{(j)}\right) \\
&= \operatorname*{argmax}_{\pi}\,(2\pi(X_i)-1)\,\Gamma_i \\
&= \operatorname*{argmax}_{\pi}\,\underbrace{(2\pi(X_i)-1)\,L_i}_{\text{Classification Objective}}\quad\underbrace{|\Gamma_i|}_{\text{Observation Weight}}
\end{aligned}
\tag{6}
$$

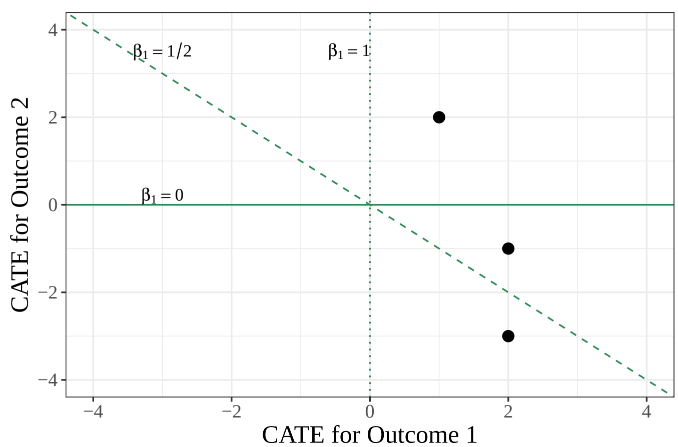
The alternative objective in the first line allows us to use simple algebraic simplifications to reach a classification objective $(2\pi(X_i) - 1) L_i$. Therefore, we show that optimizing the scalarized objective is equivalent to the weighted classification in Proposition 3 and complete the proof.

B Illustrative Examples for Algorithms

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

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

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



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

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

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

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

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

- *Case 1:* $\beta_1 = 0$. In this case, the weight for Outcome 1 is zero, and the weight for Outcome 2 is one. We first calculate Γ in this case, which consists of two parts: (1) a difference between the propensity weights $W_i/e(X_i) - (1 - W_i)/(1 - e(X_i))$, and (2) a weighted outcome $\sum_{j=1}^{K} \beta_j Y_i^{(j)}$. For data points assigned to the treatment, the difference between the propensity weights is $1/(0.5) - 0/(0.5) = 2$, and for data points assigned to the control, this difference is $0/(0.5) - (1 - 0)/(0.5) = -2$.

$$\Gamma = (2 \times 3, -2 \times 1, 2 \times -2) = (6, -2, -4) \tag{7}$$

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

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

$$\Gamma = \left(2 \times \left(\frac{1}{2}2 + \frac{1}{2}3\right), -2 \times \left(\frac{1}{2}1 + \frac{1}{2}1\right), 2 \times \left(\frac{1}{2}3 + \frac{1}{2}(-2)\right)\right) = (5, -4, 1) \tag{8}$$

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

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

$$\Gamma = (2 \times 2, -2 \times 1, 2 \times 3) = (2, -2, 6) \tag{9}$$

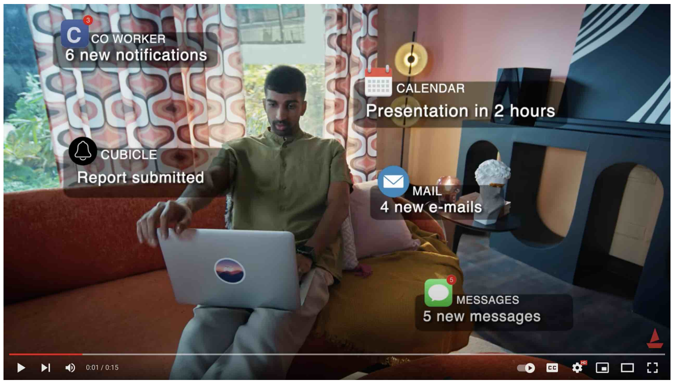
  The labels will be $(1, -1, 1)$, similar to the previous case, but the weights for our classification task are $(2, 2, 6)$.

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

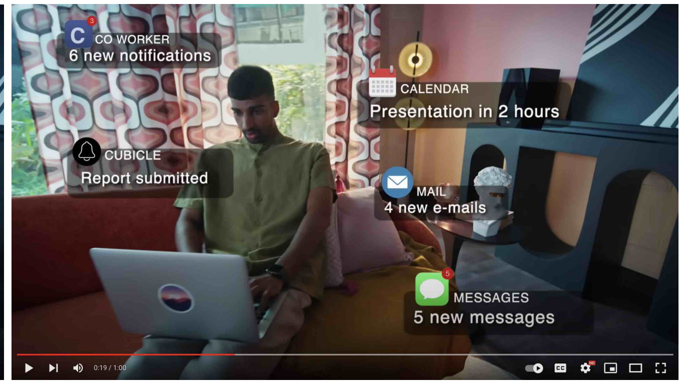
C Description of boAt Ads in the Experiment

In this section, we present supplementary information about the advertised brand and product. boAt is an audio and wearables brand that offers different wireless earphones, earbuds, headphones, smartwatches, and home audio. The advertised product, Watch Xtend, is a smartwatch with a built-in Alexa voice assistant. The ad content in both conditions presents a series of takes where different characters use the Alexa

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



(a) Non-Skippable/Short



(b) Skippable/Long

voice assistant in different contexts, highlighting the utilitarian and lifestyle aspects of the product. The primary objective of the ad campaign is to generate greater awareness about the product and brand. As such, although the ad is clickable, it is not a performance ad with a clear call for action. A click only takes users to the product's website for more information. Figure A2 shows a snapshot of different parts of both ad versions.

D Randomization Check

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

- Hypothesis Testing Approach: Our first approach uses the hypothesis testing framework. For any subcategory s, let $\pi_{s,1}$ and $\pi_{s,0}$ denote the proportion of observations belonging to the subcategory in the Skippable/Long and Non-Skippable/Short conditions, respectively. If randomization has been done correctly, we will fail to reject the following null hypothesis: $H_0 : \pi_{s,0} = \pi_{s,1}$. First, we conduct Fisher's exact test for each subcategory. Since we run multiple hypotheses, we expect a fraction of them to be significant even if the null hypothesis is true. Of 837 separate tests conducted, only 8 rejected the null hypothesis. After adjusting for multiple hypothesis testing using the Benjamini-Hochberg approach (Benjamini and Hochberg 1995), no adjusted p-value was below 0.05.[1] In Table A1, we present the proportions under Skippable/Long and Non-Skippable/Short for the top subcategories identified in Table 1, as well as the Fisher and Z-test p-values for their corresponding hypothesis tests. As shown in this table, all the p-values are greater than 0.05, indicating that we fail to reject the null hypothesis that

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

Table A1. Randomization Checks for Top Subcategories

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

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

the proportions are the same under Skippable/Long and Non-Skippable/Short conditions. This evidence suggests that randomization has been implemented correctly in our data.

- Standardized Bias Approach: In our second approach, we use the measure of *Standardized Bias (SB)*, which is commonly used in the literature to assess covariate balance. Standardized Bias is equal to the absolute difference between the means of two groups divided by the standard deviation of the covariate for the pooled sample. The common norm in the literature is to consider a Standardized Bias below 0.2 or 0.1 as evidence for covariate balance (McCaffrey et al. 2013). In our setting, we find that the maximum Standardized Bias was 0.026, which indicates that we have a covariate balance for all the pre-treatment covariates using this approach. In Table A1, we present the Standardized Bias for the top subcategories. As shown in the table, all values are substantially lower than the acceptable thresholds.
- Regression Approach: We use a regression approach to regress the treatment assignment on all the pre-treatment variables. If randomization has been done correctly, the pre-treatment variables will have no predictive power in explaining the treatment assignment. We can statistically test that by using the F-test of the regression model. We find that the F-statistic is equal to 1.02 with a p-value of 0.32, which indicates that the pre-treatment variables have no predictive power in predicting the treatment assignment and provides evidence for the validity of randomization in our study.

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

Table A2. Description of Outcome Variables.

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

data with the observations that were dropped in our sample. We perform the same analysis for the original sample and arrive at the same insight that randomization has been correctly performed.

E Supplementary Material for Experiment Analysis

E.1 Outcome Variables

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

E.2 Interpretation of Average Treatment Effect

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

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

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

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

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

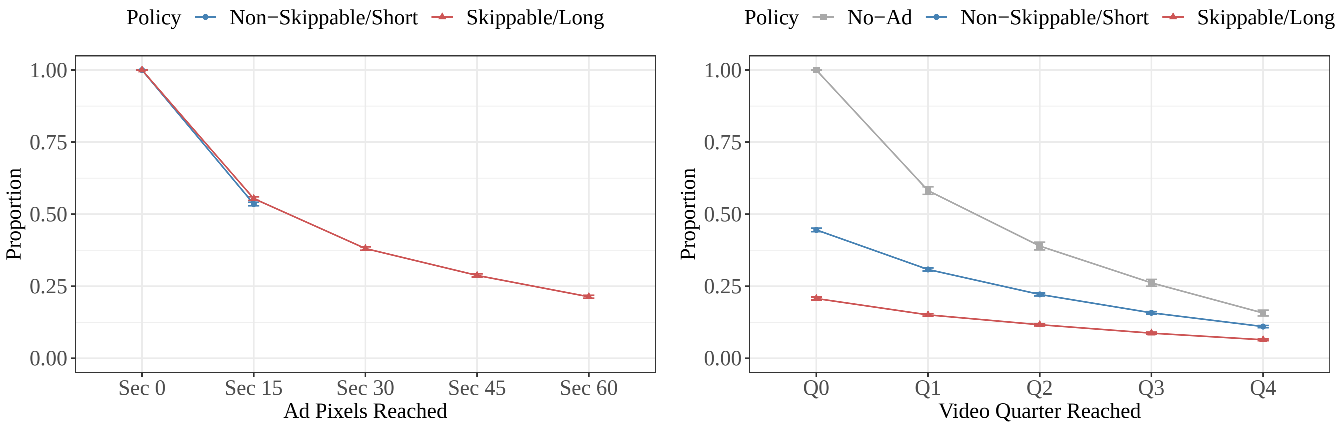
Video Consumption: Since we work with pre-roll ads, we expect the ad format to affect Video Consumption. As indicated in Table 2, we find that users in the Skippable/Long condition consume 0.38 quarters less than those in the Non-Skippable/Short condition. This is equivalent to a $0.38 \times 0.25 \times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

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

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

We now visually summarize all the findings in Figure A3. In Figure A3a, we show the proportion of users who reach a certain pixel for both Skippable/Long and Non-Skippable short ad formats. We notice

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



(a) Ad-Related Outcomes

(b) Video Related Outcomes

the higher survival rate at the completion of the Second 15 pixel for the Skippable/Long condition. This figure also shows a decay rate at the subsequent pixels for the Skippable/Long ad condition.

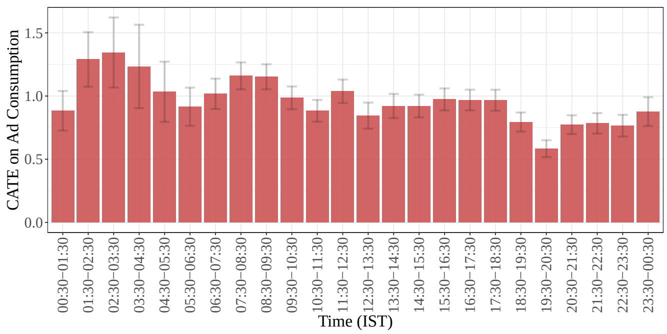
Next, in Figure A3b, we visualize the proportion of users who reach a certain video pixel for Skippable/Long and Non-Skippable/Short ad conditions, as well as these values for the No-Ad condition. Please note that we cannot show the ad-related outcomes for the No-Ad condition, as there is no ad shown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

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

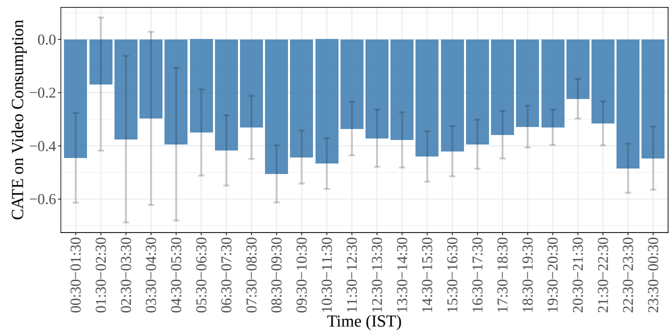
In this section, we visualize the heterogeneous treatment effects across pre-treatment covariates. In particular, we focus on (1) Hour of Day, (2) City, and (3) Operating Systems, and quantify the treatment effect conditional on different subcategories in these pre-treatment variables. Because we have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

We present the results in Figure A4. Although we find heterogeneity in treatment effects across subcategories, the substitution pattern persists across all subcategories. We now present some speculative

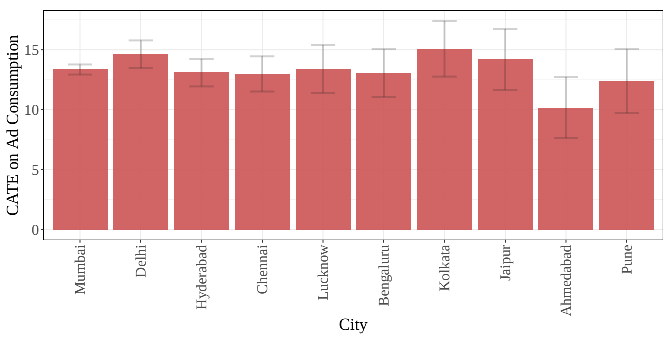
Figure A4. Heterogeneity in Treatment Effects on Both Outcomes Across Pre-Treatment Variables



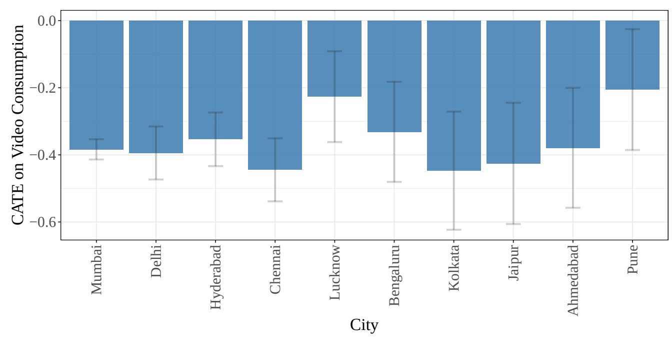
(a) Outcome: Ad Consumption, Covariate: Time



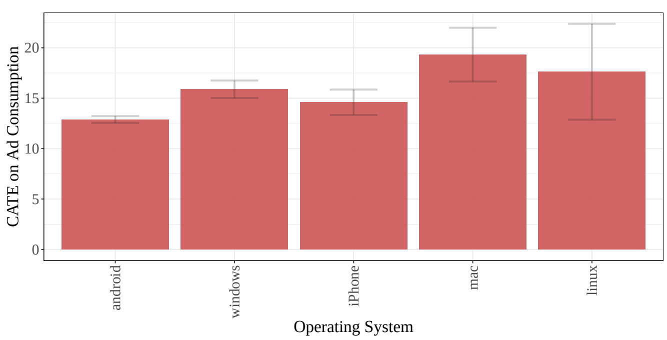
(b) Outcome: Video Consumption, Covariate: Time



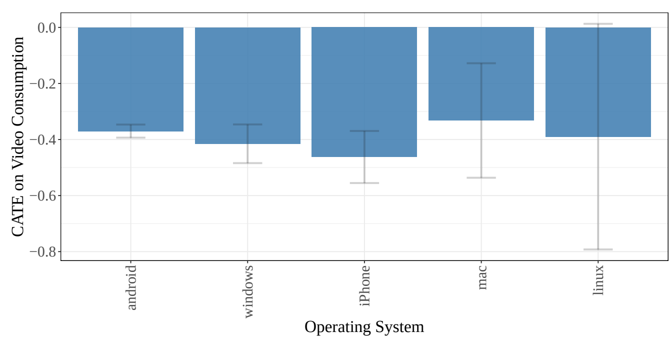
(c) Outcome: Ad Consumption, Covariate: City



(d) Outcome: Video Consumption, Covariate: City



(e) Outcome: Ad Consumption, Covariate: OS



(f) Outcome: Ad Consumption, Covariate: OS

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

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

F Details of Policies Defined in §5.2

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

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

$$\pi_\beta^{\text{MO-SCE}}(X_i) = \mathbb{1}\left(\beta \hat{\tau}_A^{\text{Train}}(X_i) + (1-\beta)\hat{\tau}_V^{\text{Train}}(X_i) \geq 0\right), \tag{10}$$

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

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

$$\Gamma_i^{(\beta)} = \left(\frac{W_i}{e(X_i)} - \frac{1-W_i}{1-e(X_i)}\right)\left(\beta Y_i^{(A)} + (1-\beta) Y_i^{(V)}\right) \tag{11}$$

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

$$\pi_\beta^{\text{MO-SCE}}(X_i) = \hat{f}_{\text{XGB}}^{(\beta)}\left(X_i, \hat{\tau}_A^{\text{Train}}(X_i), \hat{\tau}_V^{\text{Train}}(X_i)\right) \tag{12}$$

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

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

$$\pi_\alpha^{\text{SO-Mix}} = \mathbb{1}(Z_i \leq \alpha)\pi^{\text{SO-AC}}(X_i) + \mathbb{1}(Z_i > \alpha)\pi^{\text{SO-VC}}(X_i), \tag{13}$$

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

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

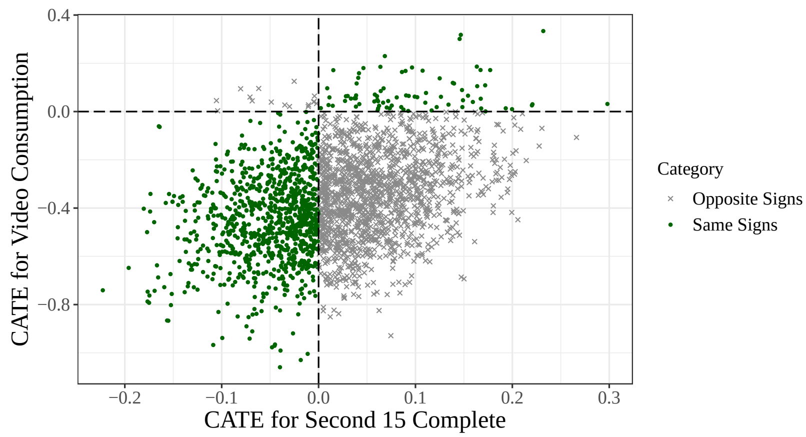
$$\pi_\alpha^{\text{Random}} = \mathbb{1}(Z_i \leq \alpha)1 + \mathbb{1}(Z_i > \alpha)0, \tag{14}$$

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

G Multi-Objective Personalization with Different Outcomes

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

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



objective for platforms is to maximize the ad revenue by having more people reach the cutoff point while keeping Video Consumption high.

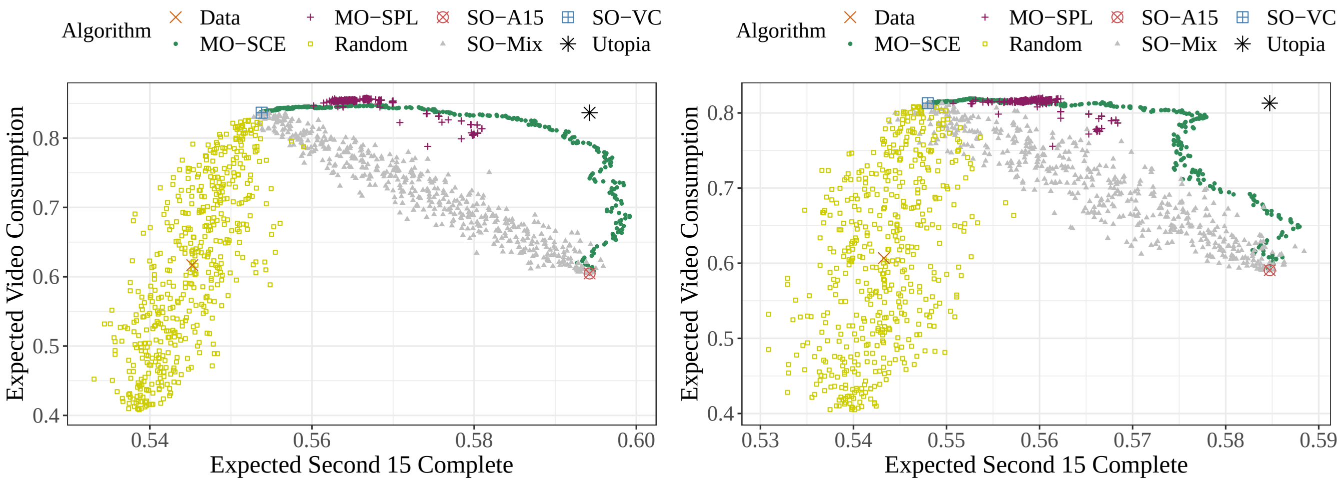
In Table 2, we presented the average treatment effect on Second 15 Complete as the outcome variable. We showed that users in the Skippable/Long ad condition are more likely to reach the 15^{th} second of the ad. However, the magnitude of the treatment effect is smaller compared to the effects on Ad Consumption.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

We first estimate the CATE on Second 15 Complete using Causal Forests, with a 10-fold cross-validation. We plot the CATE on Video Consumption against CATE on Second 15 Complete for a random sample of observations in our and present the results in Figure A5 to see the extent to which the two outcomes are in conflict. Unlike the case with Ad Consumption and Video Consumption, we note that the sign of the CATE estimates is the same for a large portion of observations. Over 41% of all observations have the same signs of CATE estimates, which means that the treatment assignment for these observations is clear: units with positive CATE estimates on both outcomes receive the Skippable/Long ad, whereas units with negative CATE estimates on both outcomes receive Non-Skippable/Short ad. Further, we find a positive correlation of 0.38 between CATE on Second 15 Complete and CATE on Video Consumption. This confirms our initial intuition that multi-objective personalization would be valuable in this setting.

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

We present the performance of all these policies using the IPS estimator when evaluated on the training and test data in Figure A6. A few insights immediately emerge from this figure. First, we note that the `MO-SCE` is more stable and covers a broader range of policies. Second, we find that the performance of `MO-SCE` is remarkably well on the training data as the Pareto frontier almost approaches the Utopia point.

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



(a) Training Data for Evaluation

(b) Test Data for Evaluation

We then turn to our *Covered Area Proportion (CAP)* measure proposed in §3.3 to evaluate the performance of the identified sets of policies. We find that the set of MO−SCE policies achieve 75% and 58% in *CAP* measure on the training and test data, respectively. The performance is worse for the set of MO−SPL policies because their coverage is more sparse: the set of MO−SPL policies achieve 51% and 43% in *CAP* measure on the training and test data, respectively.

Third, we find policies on the identified Pareto frontier with a reasonable performance relative to both single-objective policies. In our case in the main text with Video Consumption and Ad Consumption as outcomes, the extensive substitution between the two outcomes prevented us from finding a single policy on the Pareto frontier that compares well with both single-objective policies. In the multi-objective personalization problem with Second 15 Complete and Video Consumption as outcomes, we find policy $\pi_{0.760}^{\text{MO-SCE}}$ from the set of MO−SCE policies that achieve great performance when compared to either single-objective policy. Compared to the Single-Objective Second 15 Complete (SOA15) policy, this multi-objective personalized policy improves Video Consumption by 36.4%, while reducing the Second 15 Complete rate by 1.1% on the training data. On the test data, the gain in Video Consumption is 35.1%, at the same loss of 1.2% of the Second 15 Complete rate. Interestingly, compared to the Single-Objective Video Consumption policy (SO−VC), the very same policy $\pi_{0.760}^{\text{MO-SCE}}$ increases the Second 15 Complete rate by 6.1% relative to the SOVC policy, while reducing Video Consumption by 1.4%. On the test data, the gain in the Second 15 Complete rate is 5.4%, and the drop in Video Consumption is 1.9%. Together, our results show that the multi-objective personalization framework can be applied to a variety of settings and generate gains beyond the single-objective personalization framework.

References

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

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

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

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

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

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

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