Personalized discount targeting with causal machine learning

Short Paper

Alex P. Miller, Kartik Hosanagar
The Wharton School, University of Pennsylvania
alexmill@wharton.upenn.edu, kartikh@wharton.upenn.edu

Abstract

The low cost of digital experimentation and increasing capabilities of machine learning algorithms have opened up new avenues for personalization in online retail. In this project, we describe how firms can combine these techniques to find an optimal targeted discount strategy. We cast this task within a simple decision-theoretic framework, solve for the optimal solution, and describe how to use causal machine learning methods and data from an online experiment to estimate the requisite model parameters. To validate our methodology empirically, we apply it to data from a randomized experiment at an online retailer. We demonstrate that our proposed targeted discount strategy can be estimated using real-world data with sufficient accuracy to result in increased profits. By accounting for the discount rate and heterogeneity in both baseline response rates and treatment effects, our proposal significantly outperforms existing price-agnostic targeting practices.

Keywords: personalization, targeting, A/B testing, machine learning, discounting, e-commerce

Introduction

In this project, we demonstrate a significant amount of value can be realized by ensuring that personalization techniques adapted from statistics and machine learning are suited to maximize appropriate economic objectives. We study this phenomenon in the context of discount targeting, for which we propose a novel and flexible methodology for determining optimal targeting policies. Our approach combines a decision theoretic framework with modern techniques from machine learning that use experimental data to estimate conditional treatment effects. Using our framework, we demonstrate how the optimal discount strategy depends on targeting customers based on a calibrated trade-off between their individual baseline purchase rate and heterogeneity in their response to the discount. This is different from literature on targeting in non-discount settings which previously focused on targeted strategies based on only one of these quantities at a time. We assess the real-world value of our theoretical methods using data from an experiment at an online apparel retailer. We demonstrate that our proposed targeted discount strategy can be estimated using data from an A/B test with sufficient accuracy to be profitable in this empirical context and show that our proposed technique, by accounting for heterogeneity in both baseline response rates and treatment effects, outperforms other competitive targeting policies.

Background & related research

Promotional price discounts have been used by firms in a variety of settings for decades to strategically engage their customers and increase sales (Bawa and Shoemaker, 1987, Goodman and Moody, 1970). As the ability emerged to use discounts an effective means of selective price discrimination, firms in many industries started to increasingly *target* their discounts at the segment or individual customer level (Blattberg et al., 1995, Narasimhan, 1984). While early empirical research on the subject of targeted price discounts attempted to identify characteristics that distinguished customers that redeemed coupons from those that did not (Reibstein and Traver, 1982), later work recognized the importance of measuring the effects of discounts based on their ability to generate incremental sales (Bawa and Shoemaker, 1989). Later still, research started to address the problem of identifying an optimal set of customers to target with promotional discounts (Johnson et al., 2013, Rossi et al., 1996). However, in these studies researchers assume that retailers have access to a panel of customer purchase history data; as such targeting is limited to only a subset of existing customers and to those firms with a large historical database of price variation and sales data on a stable assortment of products and brands.

While such strategies may be suitable for traditional brick and mortar retailers or grocery chains, the online retail environment presents a unique set of challenges that limits the value of the aforementioned research. For one, the cost of changing product or service offerings is much lower for e-commerce and software firms. This allows for rapid and continuous iteration on many dimensions, which means yesterday's sales data may have little value for informing today's strategy. And not only is it common for supply-side factors to change rapidly, but the mix of users visiting an online storefront can vary dramatically over time as well. Changes in online advertising strategy or search engine rankings allow for firms to quickly reach entirely new segments of customers with behavioral profiles that may differ from previous cohorts. However, these challenges also come with new opportunities. Whereas traditional retailers could only effectively target customers with a known mailing address, online retailers have the ability to use technographic characteristics and trace geographic signatures to distinguish between every user on their website—whether or not they have an online account or have made a previous purchase. And, though the online retail environment can change rapidly and continuously, e-commerce firms have an unprecedented ability to conduct controlled experiments. The use of so-called A/B tests—which can be deployed both freely and instantaneously—give firms the opportunity measure the causal effects of their online strategies and keep pace with the changing digital landscape.

As these new testing and targeting technologies developed over the last two decades, statistical methods for exploiting these capabilities have also emerged. Recent advances in causal inference have demonstrated that machine learning can be used to obtain suitably flexible estimates of treatment heterogeneity in high dimensional settings from experimental data in a variety of domains (Chernozhukov et al., 2018, McFowland III et al., 2018, Wager and Athey, 2017). Further, many online services have emerged allowing firms of all sizes to apply these sophisticated methodologies for efficiently segmenting their online user base along a number of characteristics (Agarwal et al., 2016, Navot, 2018). However, such methods are typically based on general-purpose algorithms that do not take into account the revenue and cost implications of an individual targeting campaign. In this project, we demonstrate that in the context of *discount* targeting, such general-purpose methods result in sub-optimal outcomes for profit-maximizing firms. We describe a framework for estimating the optimal targeted discount policy using data from online experiments that can be evaluated quickly and directly on the population of interest. As such, our approach is suitable for deploying and targeting *ad hoc* marketing campaigns in quickly evolving e-commerce environments.

Decision-theoretic model for optimal discount targeting

We cast the task of determining an optimal discount strategy as a formal statistical decision problem (Berger, 1980). We derive the optimal policy for targeted discounts in a generic setting with multiple treatment arms at various discount levels, and then focus the remainder of our analysis on the simpler two-arm scenario. While our approach has potential value in other contexts, we focus mainly on aspects of the problem relevant to a generic e-commerce retailer.

Problem set up

Consider an e-commerce firm that observes a continuous stream of users to their online storefront. When a user (indexed by i) arrives, the firm observes $X_i \in \mathcal{X}$, a K-dimensional vector of observables, and must decide instantaneously on a treatment $T_i \in \{0, \dots, J\}$ to which the user will be assigned. Each treatment T has an associated price discount $d_T \in [0,1]$ which can be offered to the user in a standardized way. (In the e-commerce setting, one can think of each treatment intervention as a banner at the top of the retailer's website advertising a $d_T \times 100\%$ discount.) In our framework, the treatment $T = 0$ will canonically serve as the "control" or *status quo* treatment of offering no additional discount ($d_0 = 0$) and all treatments for $T \geq 1$ will have a strictly positive discount ($d_T > 0$). For each user, the firm observes whether or not their session ends with a purchase, indicated by outcome variable $Y_i \in \{0,1\}$. We denote the conditional response function

$$\mu_t(x) := \mathbf{E}\left[Y_i \mid T_i = t, X_i = x\right] = \Pr[Y_i = 1 \mid T_i = t, X_i = x],$$

which represents the conditional probability that a user with observed covariate $X_i = x$ will convert ($Y_i = 1$) when assigned treatment $T_i = t$. Lastly, assume the firm observes marginal revenue r and marginal cost c for each purchase on their site. The firm's observed profit for a given user can then be expressed as the following random variable:

$$\pi_i = Y_i(r(1 - d_{T_i}) - c)$$

This formula captures whether a user converts (Y) multiplied by the marginal profit observed when a discount d_T is applied to the full-price revenue r, less marginal costs c.

We can now define the firm's objective function, conditional on treatment assignment t, as the expected profit for a given user with covariate vector x:

$$\begin{aligned}\Pi(t \mid x) &:= \mathbf{E}[\pi_i \mid T_i = t, X_i = x] \\ &= \mu_t(x)(r(1 - d_t) - c)\end{aligned}$$

Optimal targeting policy

It is clear from the problem setup described above, in which a firm must assign one treatment arm and faces no budgetary constraints, that the optimal strategy is to assign each customer to the treatment t that has the highest expected profit $\mu_t(x)(r(1 - d_t) - c)$. Formally, if we let $\delta(x) \in \{0, \dots, J\}$ represent the treatment assignment for a customer with covariate x, we can express the optimal assignment rule δ^* as:

$$\begin{aligned}\delta^*(x) &= \arg\max_t \Pi(t \mid x) \\ &= \arg\max_t \mu_t(x)(r(1 - d_t) - c)\end{aligned}$$

In cases where a firm does face a budget constraint, such as in direct mail where the implementation of an intervention is itself costly, the firm will face a knapsack-like constrained optimization problem in deciding whom to target (Imai and Strauss, 2011, Johnson et al., 2013). However, in the e-commerce setting, the lack of a budget constraint is justified if we assume inventory has already been purchased and any incurred marginal costs (such as shipping) will be covered by the marginal revenue generated by each purchase. Further, in the context considered here, the intervention is conditional on a user arriving to the firm's website, at which point the firm faces no cost for advertising a discount.

For the remainder of the paper, we simplify the problem by considering a situation in which we wish to find the optimal targeting strategy for a single proposed discount intervention of $d \times 100\%$ ($T_i = 1$), compared to a control or *status quo* intervention of offering no discount at all ($T_i = 0$). Using the optimality criterion above, it is optimal to offer a user a discount ($\delta^*(x) = 1$) if and only if the following condition is met:

$$\mu_1(x)(r(1 - d) - c) > \mu_0(x)(r - c) \tag{1}$$

To develop better intuition about what this condition means in concrete terms, it will be useful to define conditional average treatment effect (CATE) of the treatment $T_i = 1$ for a user with $X_i = x$, relative to control condition:

$$\tau(x) := \mu_1(x) - \mu_0(x) = \mathbf{E}[Y_i \mid T_i = 1, X_i = x] - \mathbf{E}[Y_i \mid T_i = 0, X_i = x] \tag{2}$$

With suitable algebraic rearrangements, Eqns. (1) and (2) can be combined to arrive at the following formula for the optimal criterion for a targeted discount intervention:

$$\tau(x)(1 - d - c/r) > \mu_0(x)d \tag{3}$$

In this form, we can explain in much more intuitive terms precisely what the targeting criterion is accomplishing. In particular, notice that the term $1 - d - c/r$ on the left hand side represents the firm's final margin on a purchase made with a discount. Also, we can view the expression on the right hand side as the expected cost of offering a customer a discount when they would have made a purchase without the incentive. In this case, the firm is essentially giving away $d \times 100\%$ of their margin without any added benefit. Combining these definitions together, we can describe the optimal targeting rule in the following way: a customer should be offered a discount only if doing so increases the their margin-weighted purchase probability ($\tau(x)(1-d-c/r)$) by more than the expected cost of offering them a discount when they would have made a purchase without it ($\mu_0(x)d$).

Furthermore, when expressed in the form found in Eq. (3), it is immediately obvious that the optimal targeting condition requires individual-level estimates of *both* treatment heterogeneity and baseline response rates. This is in contrast to existing research on target marketing in the context of retention and direct mail campaigns. Research in these areas has highlighted the merits of targeting customers least likely to take a desired action (e.g., customers who may be at the lowest risk of contract renewal) (Ascarza and Hardie, 2013, Neslin et al., 2006) or targeting all customers that will respond positively to a given marketing campaign (Ascarza, 2018, Radcliffe, 2007). However, our finding demonstrates that—when a marketing campaign offers a promotional price incentive—the optimal targeting policy must consider *both* baseline response rates and responses to treatment. The key distinction highlighted by our model is that the discount induces a non-zero cost of targeting customers who would already take action without the intervention ($\mu_0(x)d$ in Eq. 3).

Experimentation & estimation framework

While we have established the relevant theoretical foundations for optimal targeting in the online retail environment, we have yet to describe a comprehensive framework for how firms can implement this strategy in a feasible way. We elaborate on some of the practical details of this methodology in our empirical application, but here we lay out our targeting framework at a high level. It consists of three primary phases:

1. *Experimentation:* The firm will choose discount level d and run an A/B test in which a randomized subset of users are assigned to the discount treatment condition. In the process, they will gather data on targetable customer features X_i, individual conversion responses Y_i, and treatment assignments T_i. Because we propose using machine learning methods to model customer heterogeneity, the set of variables a firm can include as predictors is quite flexible; in practice, these might include technographic, geographic, demographic, behavioral, purchase history, and other customer relationship data if available.
2. *Estimation:* Using the experimental data $\{X_i, Y_i, T_i\}$ gathered in the first phase, the firm can use modern machine learning techniques to estimate the conditional response and treatment effect functions: $\hat{\mu}_0(x)$ and $\hat{\tau}(x)$ (estimation of these functions is discussed in the following section). Factoring in their relevant revenue and cost parameters, the firm can use these functions to develop at a targeting policy based on the optimal strategy derived earlier.
3. *Targeting:* For customers that arrive to the website moving forward, the firm observes their covariates x, evaluates the targeting criterion $\hat{\tau(x)}(1-d-c/r) > \hat{\mu}_0(x)d$, and offers a discount if the criterion is met.

Empirical application with real-world data

Up to this point, we have derived the theoretically optimal personalized discount policy and described in broad terms how a firm might use this result in a targeted marketing campaign. However, we have yet to show that our findings have value in real-world settings where there are many reasons our theoretically-optimal strategy might fail. For example, the common sample sizes used in A/B tests and the limited number features that are observable to online firms can make it difficult to estimate the individual level response and treatment effect functions required for optimal targeting. If the estimates of $\hat{\mu}_0(x)$ and $\hat{\tau}(x)$ are too noisy, it may be more profitable to fall back on simpler targeting rules that don't require such fine-grained distinction between customers on multiple dimensions. As such, it is important to study our proposed strategy in an empirical setting with practical limitations common in real-world e-commerce environments. We use the remainder of this paper to address this topic.

Empirical context and dataset

We gathered data from a two-armed experiment at a mid-sized online apparel retailer operating in the United States. The experiment, conducted on domestic users in the US, consisted of allocating users to either a control arm in which no discount was offered or a treatment arm in which a 20% discount was clearly advertised in a large banner on the website's homepage with a smaller banner that was placed at the top of every page. Ninety percent of users were randomly allocated to the discount treatment with the remaining 10% serving as a control group. Our total sample includes 59,353 session-level observations; baseline conversion rate in the control and treatment conditions was measured at 3.9% and 5.0%, respectively.

Along with the conversion responses for each user, we observe a set of technographic characteristics that are commonly accessible to any web server. This includes a user's device type (desktop, tablet, mobile), operating system, web browser, screen dimensions, referral source, and—when the user arrived through a search engine that reports it—search query information. These are features commonly used by personalization platforms in many online settings (Hannak et al., 2014). The firm also observed each user's IP address which they map to ZIP codes using a geolocation service.[1] Because many of these features are categorical, our raw data matrix is inherently high-dimensional. Given the success of single value decomposition (SVD) as a dimensionality-reduction technique in other supervised learning tasks with high-dimensional data, we employ SVD to preprocess our covariate matrix in this application (Deerwester et al., 1990, Sarwar et al., 2000, Wall et al., 2003). In particular, we approximate the categorical features in our data with a truncated SVD of rank 10 (Hansen, 1987); this reduces the dimension of our data matrix from 166 sparse columns to 19 dense features.

Estimation of $\hat{\tau}$ and $\hat{\mu}$

In this project, we apply the method described in Chernozhukov et al. (2018) to estimate the individual-level conditional response functions, $\hat{\tau}(x)$ and $\hat{\mu}_0(x)$, from our experimental data. At a high level, the approach takes an arbitrary supervised learning algorithm (a "base learner") and uses repeated sample-splits of the experimental data to construct an ensemble across subsamples. For each sample split, the data is randomly divided into "main" and "auxiliary" samples. In each split, a base learner is trained on the main split and then used to make out-of-sample predictions on the auxiliary split; the resulting predictions are then used in a carefully constructed regression on the true outcomes of the auxiliary data. At the end, the coefficients of the regressions and the predictions at each stage are averaged over and then re-combined to generate ensembled final predictions. For full details, the reader is referred to the original paper.

Though other approaches may be suitable, this technique has several desirable properties in our context. In addition to providing a means for inference on otherwise non-parametric machine learning models, ensembling a base learner's predictions across a large number of splits very often results in models that are more robust than single classifiers trained on the entire sample. Furthermore, because the meta-model can be built on top of any base learning algorithm, it offers considerable flexibility in the modeling of potentially complex conditional response functions. Lastly, the sample-splitting procedure for estimating $\hat{\tau}(x)$ allows us to naturally and concurrently generate estimates of $\hat{\mu}_0(x)$ and $\hat{\mu}_1(x)$ with little additional computation.

In our implementation, we use a gradient boosted decision tree classifier as our base supervised learner, with parameters that were tuned using similar data, but from a different experiment by another firm. Our models are trained using all the covariate data described above, with each user's binary conversion response $Y_i \in \{0, 1\}$ as the prediction target. Our final estimators aggregate the results of 100 base learners, trained on even splits of our experimental data, stratified by treatment condition. At the end of the training period, we are able to calculate ensembled predictors for each of the functions $\hat{\tau}(x)$, $\hat{\mu}_0(x)$, and $\hat{\mu}_1(x)$.

Alternative targeting policies

Before moving on to our empirical findings regarding the performance of our proposed targeting policy, it will be instructive to identify other reasonable policies a firm might use in its place. To continue, we first define a standard notation for specifying a targeting policy with a single function. For a given customer with observable covariate x, a targeting policy will be defined by its targeting function $f(x)$ and the decision δ_f to

[1]ZIP codes were historically developed for use by the United States Postal Service to designate approximate geographic mailing areas. While they have many limitations that make them unsuitable for precise geolocation, our results demonstrate they can nonetheless be useful for online targeting.

offer a customer the discount treatment ($T = 1$) will be determined by the sign of $f(x)$:

$$\delta_f(x) = \begin{cases} 1 & \text{if } f(x) > 0 \\ 0 & \text{if } f(x) \le 0 \end{cases}$$

In this notation, the targeting function for the optimal policy implied by Eq. (3) becomes:

$$f(x) = \tau(x)(1 - d - c/r) - \mu_0(x)d$$

As a starting place, it makes sense to consider a *non-targeted* (or uniform) policy that treats all customers the same. The profit-maximizing choice for such a policy will be identical to the optimal policy above with the individual level estimates of $\tau(x)$ and $\mu_0(x)$ replaced by their population expectations. Using the overline to denote population means, ($\overline{\tau} = \mathbf{E}_x[\tau(x)]$ and $\overline{\mu_0} = \mathbf{E}_x[\mu_0(x)]$), the targeting function for this policy can be expressed as:

$$f(x) = \overline{\tau}(1 - d - c/r) - \overline{\mu_0}d$$

As for other possible targeting policies, it may prove instructive to disentangle the effects of modeling baseline response heterogeneity and treatment effect heterogeneity. For example, consider a "fixed baseline" policy that sets the baseline response to the population average, but allows for individual level treatment effects:

$$f(x) = \tau(x)(1 - d - c/r) - \overline{\mu_0}d$$

And analogously, consider a "fixed treatment effect" policy that allows for individual-level variation in baseline response but assumes a constant treatment effect across the population:

$$f(x) = \overline{\tau}(1 - d - c/r) - \mu_0(x)d$$

Comparing these policies to the optimal policy in which both treatment effects and baselines are allowed to vary will be useful for understanding how much each of these factors contribute to the overall profitability of our approach. Lastly, we consider a well-known policy that has been mentioned many times in the literature on targeted marketing (Lo, 2002, Rzepakowski and Jaroszewicz, 2012), which is to target all customers with positive treatment effect:

$$f(x) = \tau(x)$$

This strategy, which we will refer to as the "true lift" approach, will serve as a useful baseline for considering the value of our decision-theoretic approach with existing practices.

Evaluation & Empirical Results

The task of estimating the value of a proposed targeting policy maps directly on to the problem of off-policy evaluation in the literature on reinforcement learning. If a user was assigned one treatment in our experimental data and a targeting policy would have assigned them to the opposite treatment, we cannot observe their counterfactual response and thus must impute this value in a reliable way to estimate the profitability of the proposed policy. Given that we have data that were generated with random treatment arm assignment, we can use the method of inverse probability weighting (IPW), which is known to provide unbiased estimation of off-policy rewards (Dudík et al., 2014, Horvitz and Thompson, 1952).

Let $\{y_i, x_i, t_i\}$ represent the observed response, covariate, and treatment assignment data from the N users in our experiment. We can express the IPW estimator for the average profit per user of a policy with targeting function f with the following expression:

$$\text{IPW estimator: } \hat{\pi}_f = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{t_i = 1\}\frac{\mathbb{1}\{f(x_i) > 0\}}{\Pr[t_i = 1]} y_i(r - d - c) + \mathbb{1}\{t_i = 0\}\frac{\mathbb{1}\{f(x_i) \le 0\}}{\Pr[t_i = 0]} y_i(r - c)$$

In our case, because treatment assignment is independent of any other variable, a direct estimator that ignores the sampling probabilities will also be unbiased. Such an estimator can be constructed by estimating the proportion of users to be allocated to each treatment and using the average response of users who happen to be in their counterfactual treatment as the mean response of each group. We will report the results of both estimators.

Using the same data to estimate both the targeting function $\hat{f}$ and the expected profits $\hat{\pi}_f$ can overstate the effectiveness of a policy on out-of-sample data. As such, our final results are derived by averaging over 100 iterations of Monte Carlo cross-validation, where in each iteration 2/3 of our data is used for training the models needed to estimate a policy's target function $\hat{f}$ and the remaining 1/3 of the data is used to estimate the policy's profit.

Lastly, to arrive at meaningful empirical results, we must choose suitable values for the exogenous parameters in our model. The discount rate d is set to 20% based on the treatment that was deployed in our experimental data. Because we cannot directly observe the firm's exact margins in our data, we report the empirical results with the revenue set to a normalized value $r = 1$ and the cost parameter $c = 0.1$. However, we emphasize that analyses performed with other values of the cost parameter result in qualitatively similar findings as those reported below.

We performed the evaluation process described above for each of proposed targeting policies identified earlier. Our primary outcome metric is the average profit observed across Monte Carlo iterations $\mathbf{E}[\hat{\pi}]$. To facilitate comparison, we also compute the average lift or percentage gain observed for each policy compared to the profits observed for the non-targeted uniform policy. If $\hat{\pi}_0$ are the profits observed from the uniform policy, the average gain from an alternative policy f is given in percentage terms as:

$$\mathbf{E}[\hat{\Delta}_f] = \mathbf{E}[(\hat{\pi}_f - \hat{\pi}_0)/\hat{\pi}_0 \times 100]$$

In Table 1, we have summarized the empirical results of both profit and lift metrics calculated across cross validation folds, using both direct and IPW estimators. Each row in the table reports results corresponding to the aforementioned targeting policies. For all policies except the fixed treatment effect policy, the reported average lift is different from the uniform baseline with $p < 0.001$. Furthermore, in Figure 1, we have plotted the results of the IPW-estimated profit gains, calculated across Monte Carlo iterations for each targeting policy—including the raw distribution of cross validated profits estimates (top figure), as well as the distribution of sample means (bottom figure).

		Direct		IPW	
Policy name	Targeting function $f(x)$	Profit $\mathbf{E}[\hat{\pi}_f]$	Gain $\mathbf{E}[\hat{\Delta}_f]$	Profit $\mathbf{E}[\hat{\pi}_f]$	Gain $\mathbf{E}[\hat{\Delta}_f]$
Uniform	$\overline{\tau}(1-d-c/r) - \overline{\mu_0}d$	0.0353	—	0.0349	—
True lift	$\tau(x)$	0.0363	**2.7%**	0.0365	**4.5%**
Fixed treatment effect	$\overline{\tau}(1-d-c/r) - \mu_0(x)d$	0.0355	-0.4%	0.0351	-0.6%
Fixed baseline	$\tau(x)(1-d-c/r) - \overline{\mu_0}d$	0.0380	**7.6%**	0.0381	**9.1%**
Optimal policy	$\tau(x)(1-d-c/r) - \mu_0(x)d$	0.0457	**8.9%**	0.0387	**10.8%**

Table 1: Empirical results for proposed targeting policies

We remark on several aspects of these results. First, qualitative findings about the performance between different policies are essentially the same using either estimator; we will comment on the IPW results. We see that the true lift model, which offers discounts to all users with a positive estimated treatment effect, does result in a positive expected profit the uniform, non-targeted policy (+4.5%). However, comparing the true lift model to the fixed baseline model highlights the importance of choosing appropriate decision-theoretic thresholds for targeting in this setting. The only difference between these two models is that individual treatment effects are scaled by the firm's margin, and then compared against the average cost of offering a discount to users who would have converted without the incentive. Thus, without estimating any additional models, accounting for the appropriate revenue implications of a discount in a decision-theoretic framework allows us to achieve significant gains above the common true lift targeting strategy.

Further, when comparing the fixed baseline (+9.1%), fixed treatment effect (-0.6%), and optimal policies (+10.8%), it appears that most of the gains of the optimal policy are due to its ability to account for treatment effect heterogeneity, as opposed to baseline response heterogeneity. The fixed treatment effect policy, that assumes all users respond to the discount identically, but allows for their baseline response rates to

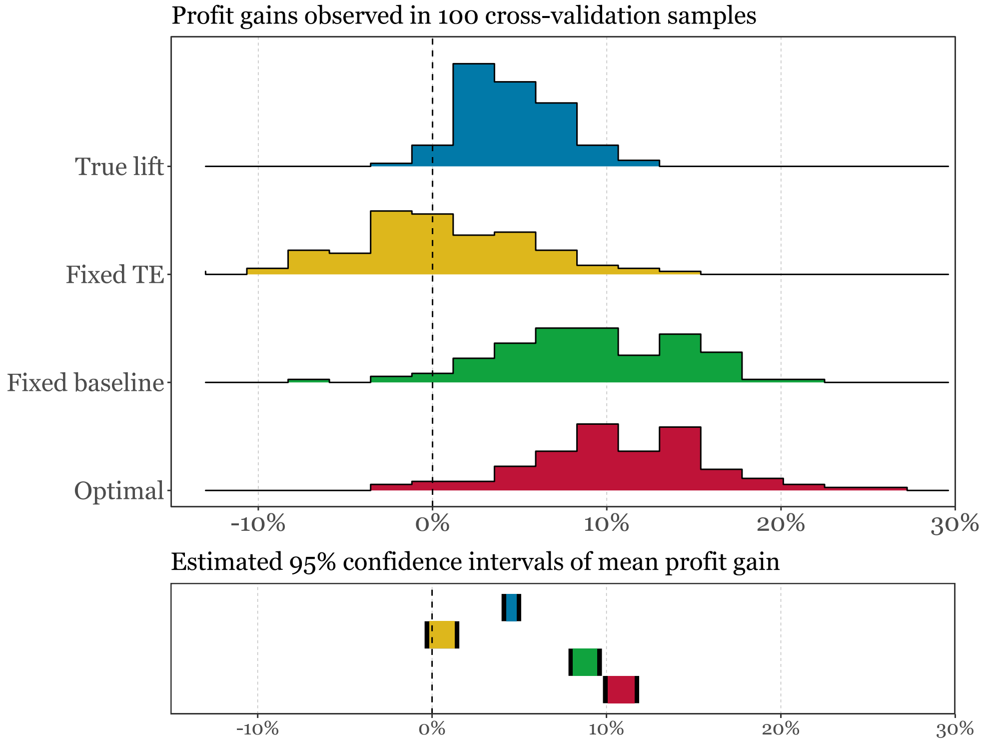

Figure 1: IPW estimates of profit gains over uniform, non-targeted policy

vary across users, fails to be a profitable strategy by itself. But by flipping these dimensions—fixing baseline response rate across users and allowing for individual-level treatment heterogeneity—results in a targeting policy that is near the best-performing strategy. However, we do observe a form of complementarity when combining *both* treatment and baseline response heterogeneity in our decision theoretic framework, demonstrating that our theoretically-derived optimal targeting strategy can be estimated with sufficient accuracy in empirical settings to yield increased profit.

Conclusion

In this project, we proposed a decision-theoretic model of targeted discounts and found the optimal solution in terms of functions that can be feasibly estimated from common experimental data. Using techniques at the intersection of causal inference and machine learning, we applied this model to experimental data and estimated that the firm could increase their profits by more than 10% over a uniform policy if they were to adopt our proposed targeting policy. We estimated this to be more than twice the gain that existing, price-agnostic targeting practices can provide. The model introduced here naturally lends itself to extension and in future work we hope to develop the model to allow for more flexibility in the cost and revenue parameters and to investigate its value in other empirical contexts. However, even in its current form, we believe the methods described in this project represent a valuable innovation in the practice of online personalization.

References

Agarwal, A., Bird, S., Cozowicz, M., Hoang, L., Langford, J., Lee, S., Li, J., Melamed, D., Oshri, G., Ribas, O., Sen, S., and Slivkins, A. (2016). A Multiworld Testing Decision Service. *ArXiv preprint*.

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

Ascarza, E. and Hardie, B. G. (2013). A joint model of usage and churn in contractual settings. *Marketing Science*, 32(4):570–590.

Bawa, K. and Shoemaker, R. W. (1987). The effects of a direct mail coupon on brand choice behavior. *Journal of Marketing Research*, 24(4):370–376.

Bawa, K. and Shoemaker, R. W. (1989). Analyzing incremental sales from a direct mail coupon promotion. *Journal of Marketing*, 53(3):66–78.

Berger, J. (1980). *Statistical Decision Theory: Foundations, Concepts, and Methods*. Springer Series in Statistics.

Blattberg, R. C., Briesch, R., and Fox, E. J. (1995). How promotions work. *Marketing science*, 14(3_supplement):G122–G132.

Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I. (2018). Generic machine learning inference on heterogenous treatment effects in randomized experiments. Technical report, National Bureau of Economic Research.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6):391–407.

Dudík, M., Erhan, D., Langford, J., Li, L., et al. (2014). Doubly robust policy evaluation and optimization. *Statistical Science*, 29(4):485–511.

Goodman, D. A. and Moody, K. W. (1970). Determining optimum price promotion quantities. *Journal of Marketing*, 34(4):31–39.

Hannak, A., Soeller, G., Lazer, D., Mislove, A., and Wilson, C. (2014). Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 conference on internet measurement conference*, pages 305–318. ACM.

Hansen, P. C. (1987). The truncatedsvd as a method for regularization. *BIT Numerical Mathematics*, 27(4):534–553.

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.

Imai, K. and Strauss, A. (2011). Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign. *Political Analysis*, 19(1):1–19.

Johnson, J., Tellis, G. J., and Ip, E. H. (2013). To whom, when, and how much to discount? A constrained optimization of customized temporal discounts. *Journal of Retailing*, 89(4):361–373.

Lo, V. S. (2002). The true lift model: a novel data mining approach to response modeling in database marketing. *ACM SIGKDD Explorations Newsletter*, 4(2):78–86.

McFowland III, E., Somanchi, S., and Neill, D. B. (2018). Efficient discovery of heterogeneous treatment effects in randomized experiments via anomalous pattern detection. *arXiv preprint arXiv:1803.09159*.

Narasimhan, C. (1984). A price discrimination theory of coupons. *Marketing Science*, 3(2):128–147.

Navot, Y. (2018). A Contextual-Bandit Approach to Website Optimization. `https://www.dynamicyield.com/blog/contextual-bandit-optimization/`. (Accessed on 08/28/2019).

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., and Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of marketing research*, 43(2):204–211.

Radcliffe, N. (2007). Using control groups to target on predicted lift: Building and assessing uplift model. *Direct Marketing Analytics Journal*, pages 14–21.

Reibstein, D. J. and Traver, P. A. (1982). Factors affecting coupon redemption rates. *Journal of Marketing*, 46(4):102–113.

Rossi, P. E., McCulloch, R. E., and Allenby, G. M. (1996). The value of purchase history data in target marketing. *Marketing Science*, 15(4):321–340.

Rzepakowski, P. and Jaroszewicz, S. (2012). Uplift modeling in direct marketing. *Journal of Telecommunications and Information Technology*, pages 43–50.

Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2000). Application of dimensionality reduction in recommender system-a case study. Technical report, Minnesota Univ Minneapolis Dept of Computer Science.

Wager, S. and Athey, S. (2017). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, (just-accepted).

Wall, M. E., Rechtsteiner, A., and Rocha, L. M. (2003). Singular value decomposition and principal component analysis. In *A practical approach to microarray data analysis*, pages 91–109. Springer.