Transferring Information Between Marketing Campaigns to Improve Targeting Policies

Artem Timoshenko
Kellogg School of Management

Marat Ibragimov
MIT Sloan School of Management

Duncan Simester
MIT Sloan School of Management

Jonathan Parker
MIT Sloan School of Management

Antoinette Schoar
MIT Sloan School of Management

November 2020

Transferring Information Between Marketing Campaigns to Improve Targeting Policies

Abstract

Targeting policies are typically trained using data from field experiments. Large experiments yield more training data and better targeting policies, but large experiments are costly to implement. We show how firms can train targeting policies using smaller experiments by incorporating information from past marketing campaigns. Even though the past campaigns may involve different marketing actions and different types of customers, the transferred information can enable firms to train policies that are as profitable as policies designed using much larger experiments. In practice, firms can maintain a corpus of past targeting policies, and leverage these policies in each new campaign.

We document the benefits using field experiment data from three different firms: a luxury fashion retailer, a membership wholesale club, and a financial services firm. Transferring information between marketing campaigns consistently improves the performance of targeting policies. The performance improvement depends upon the size of the focal experiment. Transferring information is most valuable when the focal experiment is large enough to identify the relevant information in the source data, but small enough that the source data provides valuable incremental information. The performance improvement also depends upon the similarity between the source and focal problems. For firms with many past campaigns, we propose a measure that can help to prescreen potential source campaigns.

Keywords: Marketing Analytics; Targeting Models; Transfer Learning; Conditional Average Treatment Effects; Marketing Promotions; Field Experiments

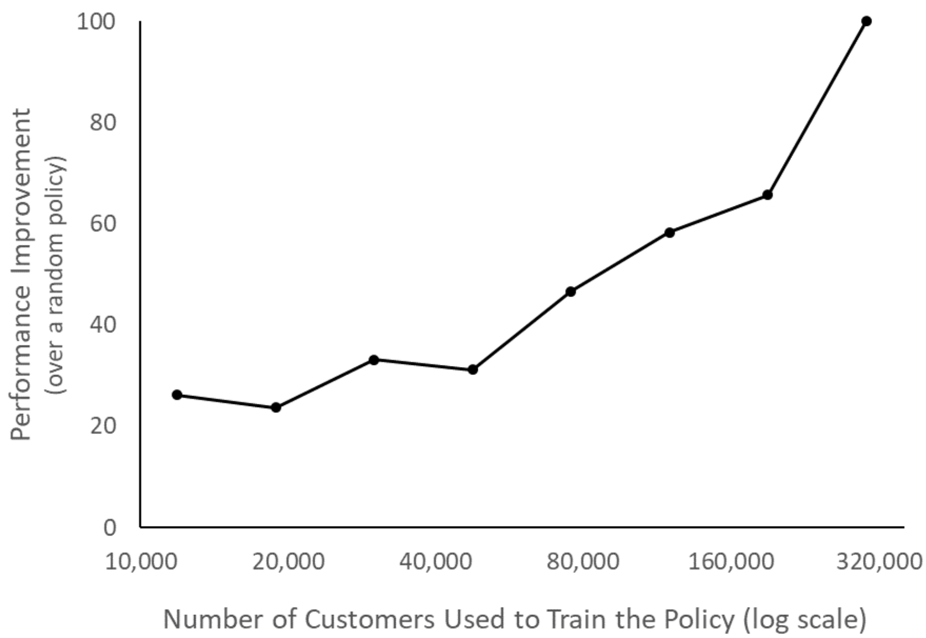
1. Introduction

Many firms want to increase the profitability of their marketing campaigns by targeting different customers with different marketing actions. Retailers send personalized coupons based on customers' purchase histories; social networks show advertisements that vary depending on users' profiles or search behavior; financial services firms offer different products according to customers' investment records. These personalized solutions are directed by targeting policies, which segment customers by their responsiveness to different marketing actions.

Targeting policies are typically trained using data from field experiments. Field experiments ensure both that there are groups of similar customers in the experimental sample who received each marketing action, and that differences in customer behavior are caused by these actions. However, when designing the experiments, firms face a tradeoff: large experiments yield more training data and better targeting policies, but large experiments can be costly (or infeasible). These costs may reflect the cost of the marketing actions themselves; firms incur costs when making outbound calls, when sending direct mail, or when purchasing digital advertisements. There may also be opportunity costs. If the firm has an existing targeting policy that is profitable, then deviating from that policy is potentially costly. The population of customers may also be limited; a Porsche dealer may have relatively few past customers that it can experiment on. All of these factors can limit the size of the initial experiments.

We illustrate the relationship between the size of the experiment and the performance of targeting policies in Figure 1. The policies in Figure 1 are trained using a direct mail experiment conducted by a luxury fashion retailer. The experiment included two conditions (*mail* and *not mail*), with customers in the treatment condition receiving a catalog containing information about the next season's products. We train targeting policies using different-sized random subsets of the data, and evaluate the average profit per customer using a holdout sample (additional details are provided in Section 4). Targeting policies trained with more data offer better performance, but obtaining large samples of experimental training data is costly.

Figure 1. Targeting Performance When the Training Data is Small vs. Large



The figure reports the performance of targeting policies trained using different-sized random samples of experimental data. The performance improvements are calculated compared to a random policy. The x-axis indicates the sample size of the training data (using a log scale). The outcome measures are indexed by setting the performance improvement to 100 when the sample size contains 300,000 observations.

We propose an approach to reduce the sample size of the experimental data required for training targeting policies. Our approach supplements the experimental data by transferring information from past marketing campaigns. These past campaigns may involve different marketing actions and different types of customers. Despite these differences, we show that the potential reductions in sample size are large. For example, the performance in Figure 1 of a policy trained using 120,000 customers can be matched using just 50,000 customers.

The insight motivating the proposed method is that the response of a customer to one marketing action may contain information about how similar customers will respond to a different marketing action. For example, if some customers always discard some types of catalogs without reading them, then the (lack of a) response to one catalog may provide information about the response to other catalogs. More generally, customers who discard catalogs may also be more likely to discards emails, and less likely to engage with other marketing communications. When this is true, we can transfer information from a past marketing campaign to help train a policy for a future campaign. We empirically demonstrate that the information transfer may be valuable even if: (a) the past campaigns involved different marketing actions, (b) the relevance of the past campaigns to the current problem is uncertain, and (c) the data from these past campaigns is no longer available (as long as we can still observe the resulting policies).

While our analysis focuses on direct mail campaigns sent by a luxury fashion retailer, we replicate our results using data from two additional firms. The first firm is a membership warehouse club. We study customer acquisition campaigns, in which the firm sent different promotions to randomly selected samples of prospective customers. The second firm is a large financial services firm that randomized direct mail and email campaigns to existing customers. These marketing communications were designed to inform the customers about the firm's products and services. Together the three firms provide variation in industry (luxury goods, wholesale club and financial services), types of customers (existing versus prospects), value proposition (promotions versus information), and marketing channels (email versus direct mail). We document consistent evidence that augmenting training data by transferring information from different source problems can significantly improve the performance of targeting models.

The performance improvement from transferring information varies, depending upon the amount of focal data available. This relationship follows an inverted U-shaped curve. When there is only a small amount of focal data, transferring information from source problems is unlikely to be effective. With very little focal data, the focal information is not sufficient to identify which variation in the source data is relevant. In contrast, if the focal data is large, then the source data provides little incremental information. Transferring information between marketing problems is most valuable when the size of the focal experiment is large enough to identify the relevant information in the source data, but small enough that the source data provides valuable additional information.

Firms often have hundreds of past marketing campaigns that could serve as a source campaign for transferring information. This introduces a new problem; how to choose a source campaign? The challenge is particularly relevant when the focal experiment is small, because there may be too little focal data to accurately weight the information in a large set of source datasets. One approach is to match the characteristics of the source campaigns with the characteristics of the focal campaign. For example, consider a focal campaign in which, at the end of February, the firm will mail a postcard offering a 25% discount on footwear, to customers who purchased footwear in the prior 12 months. We might be tempted to choose source campaigns based upon the marketing channel (direct mail), the type of mailing (a postcard), the mailing date (end of February), the value proposition (a discount), the type of discount (25% off), the product category (footwear), and/or the qualifying customers (purchasers of footwear in the previous 12 months). However, in practice, there are a large number of possible characteristics to choose from, and it is not obvious which characteristics are most informative about the relevance of the source campaign to the focal campaign.

Instead, we propose a single measure that can be used to prescreen potential source campaigns. The measure evaluates targeting policies trained using the source dataset by estimating the expected profit when these policies are applied directly to the focal problem (without tuning the policy using any focal data). The expected profit from directly applying the source targeting policy indicates the incremental information in the source campaign that is relevant to the focal targeting problem. We show that this proposed heuristic can be used to predict which source campaign will contribute larger improvements when transferred to the focal problem. We present evidence that this measure provides a better signal identifying relevant source campaigns than simply comparing the attributes of the source and focal campaigns.

In practice, we anticipate three types of applications. First, the proposed approach can improve existing targeting policies by supplementing the focal training data with information from previous marketing campaigns. We focus on this application in our empirical analysis.

Second, some past campaigns include small random samples of customers that did not receive the marketing action. For example, the luxury fashion retailers in our study has a policy of not mailing to 10% of customers in each of its direct mail campaigns. These holdout samples are designed to measure the average performance of the marketing campaigns, but are often too small to train targeting models. However, by transferring information from other past campaigns, the small holdout samples may become sufficient. In these applications, the method allows the firm to train targeting policies using existing data, without conducting a new experiment.

The third type of application involves targeting a new marketing action. If the firm wants to target an action that was never implemented in the past (e.g., a new type of promotional offer), then the firm will often need to start with a new experiment. By transferring information from past campaigns with different marketing actions, the firm may be able to use a smaller sample size in its initial experiment.

In either application, when a firm finishes a new marketing experiment, it can add that campaign to the existing corpus of source campaigns. The corpus of past campaigns provides a basis for training future targeting policies. The firm can select source campaigns from the corpus, and tune the information transfer using a small focal experiment.

While we use data from direct mail and email experiments in our empirical applications, we believe that the proposed method is both relevant and applicable to campaigns in other marketing channels. For example, a major obstacle to training targeting campaigns for in-person or telephone campaigns is the cost of conducting experiments. The proposed method may help to reduce this cost by reducing the required sample sizes. Alternatively, while conducting experiments in digital channels is often relatively inexpensive, the treatment effects tend to be smaller than in other channels, and so firms may need very large experiments to ensure there is sufficient training data (Lewis and Rao 2015). Our proposed approach can help to reduce this burden, by supplementing the training data using data from previous campaigns.

Related Literature

Our findings contribute to a growing literature in marketing focusing on the training of targeting policies. Targeting applications segment customers based upon their responsiveness to the firm's marketing actions (Guelman et al. 2012; Ascarza 2018). Responsiveness to marketing actions is a causal phenomenon, and so it is common to use field experiments to provide training data for targeting policies (Dubé and Misra, 2017; Hitsch and Misra, 2018; Chernozhukov et al. 2018; Simester et al. 2020a). The experimental data can be used to estimate conditional average treatment effects (CATEs) for each customer, and a targeting policy can assign the action with the highest predicted effect. For example, Yoganarasimhan et al. (2020) use information about customer demographics and behavior to predict the profitability of free trial promotions of different lengths. For each customer, the targeting model then recommends the promotion that is expected to yield the highest profit from that customer. An alternative approach is to estimate the optimal action without explicitly estimating CATEs. For example, Li et al. (2010) uses a contextual-bandit approach to provide personalized news recommendations, and Urban et al. (2014) and Schwartz et al. (2017) rely on multi-armed bandit methods to improve the efficiency of digital advertising.

Recent research has studied: the size of the experiments required to identify marketing fundamentals (Li et al. 2015; Lewis and Rao 2015), the exploration-exploitation tradeoff in experiments (Gittins et. al 2011; Bubeck and Cesa-Bianchi 2012), and ways to use observational data to supplement experimental data (Nichols 2007; Peysakhovich and Lada 2016). We show that data from past marketing campaigns can also be used supplement experimental data. As a result, firms may be able to train targeting policies either without conducting an initial experiment, or by greatly reducing the size of an initial experiment.

Transferring information across domains to improve the performance of machine learning applications is an active research direction in computer science, where it is typically labelled "transfer learning" or "domain adaptation" (Pan and Yang, 2010). For example, machine translation systems, voice assistants, and music recommendation systems often use word representations pretrained using large textual corpora (Zou et al. 2013; Devlin et al. 2018; Caselles-Dupré et al. 2018). Computer vision applications, such as face recognition and medical imaging, train models on large natural image datasets and then customize these pretrained models for specific applications (Xiaosong et al. 2017; Ng et al. 2015).

The idea that firms can train targeting models, by tuning a large corpus of pre-trained source information to a new campaign, is also related to recent research in natural language processing and image recognition, which rely on pretrained deep neural networks using large external datasets. The neural networks are then fine-tuned to the new task (Oquab et al. 2014; Donahue et al. 2013). These methods are also adopted in marketing applications (Liu et al. 2020; Burnap et al. 2020). Our empirical applications feature data from field experiments in non-digital business environments, which cannot leverage the benefits of pretrained neural networks due to the limited sample sizes in the source problems. However, our method relies upon pre-trained policies (or estimates of treatment effects), and transfers this information to the focal campaign. Because this information transfer requires only a parsimonious number of parameters, we can accomplish the transfer using only a small amount of focal data.

We can also compare our research to the literature on ensemble methods (Dietterich, 2000; Zhou, 2012). Ensemble methods estimate multiple predictive models and then combine predictions from these models to fit the outcome variable. For example, in a political science setting, Grimmer et al. (2017) use ensemble methods to predict how message content affects participants with varying partisan and ideological orientations. Intuitively, different machine learning methods capture different variation in the data, and a combination (ensemble) of methods can generate a better prediction. In our research, we combine targeting policies estimated on different datasets. We use measures of responsiveness to firm's actions in different marketing campaigns to improve performance in the focal targeting problem.

2. Description of the Data and the Business Problem

We focus the analysis on data provided by a major US luxury fashion retailer. At the time of study, the retailer operated physical stores in many cities, together with an online website. The assortment spanned men's and women's shoes and apparel, jewelry, accessories, and beauty products.

Our data includes complete transaction histories from January 2015 to March 2020 for both the online and in-store channels. Transactions have customer identifiers; the online and offline customer identifiers are matched (by a third-party firm) using credit card and address information.

The retailer regularly sends catalogs and other direct mail to its existing customers to announce and promote new products. The focus of this direct mail varies, and could include specific brands, specific product categories, or specific seasons. Between 2017-2019, the firm sent over 450 different direct mail campaigns to different samples of customers (at different times).

To measure the overall impact (ATE) of each direct mail campaign on incremental purchases and profits, the retailer conducted A/B tests. The firm held out a 'no mail' control sample of approximately 10% of the qualified households for each campaign. We received data describing the criteria used to select which customers were eligible to receive the mailing, and the circulation files identifying households randomly assigned to the mail and no mail conditions.

The targeting decisions for the direct mail campaigns were traditionally made using business rules. For example, the high-spending customers received a fashion catalog, and only customers with high spending in the beauty category received a beauty catalog. In 2019, the retailer began to use targeting models to make the circulation decisions for direct mail. The firm wanted to leverage past experimental data from the A/B tests to target customers based on responsiveness to marketing promotions. However, many of the past campaigns had too few customers in their holdout conditions to train new targeting policies.

We propose supplementing the information in the focal training data with information from other marketing campaigns. The data from different marketing campaigns represent potential sources of information about customer responsiveness to marketing communications. For example, information from a past Fashion catalog may help to improve the targeting of a future Beauty catalog. We describe how to transfer information between campaigns in the next section, and apply the proposed method to the empirical data in Sections 4 and 5. To demonstrate robustness, we also replicate our primary findings in Section 6 using data from two additional firms: a membership wholesale club and a financial services firm. We relegate descriptions of these datasets to the Appendix.

3. Transferring Information from Source Problems

Problem Formulation

We consider a firm that chooses which marketing action each customer should receive. We assume a sample of customers $\mathcal{I}$, and a finite action space A. Each customer $i \in \mathcal{I}$ is characterized by a vector of targeting variables X_i. The firm wants to train a targeting policy $\mathcal{P}$ mapping targeting characteristics to marketing actions, i.e. $t_i = \mathcal{P}(X_i),\ t_i \in A$. For each customer i and marketing action $a \in A$, we assume a monetary outcome $Y_i(a)$ if customer i is treated with a marketing action a. The value of the targeting policy $\mathcal{P}$ can then be defined as follows:

$$V(\mathcal{P}) = \frac{1}{|\mathcal{I}|}\sum_{i\in\mathcal{I}} Y_i(t_i)$$

To illustrate the notations, consider a problem faced by the retailer in our study. The retailer may want to decide which customers should receive a Fall Holiday catalog. If there is only a single version of the catalog, we can write the firm's action space $A = \{no\ mail, mail\} = \{0,1\}$. For each customer, the targeting variables X_i includes variables describing customer-specific purchase histories in different product categories, together with demographic indicators. The outcomes under the two mailing conditions, $Y_i(0)$ and $Y_i(1)$, indicate the profits contributed by customers in each condition, adjusted for the mailing cost of the catalog. The goal is to maximize expected profit by designing a targeting policy that uses targeting variables X_i to separately recommend whether to *mail* ($t_i = 1$) or to *not mail* ($t_i = 0$) a catalog to each customer.

We assume two inputs that the firm can use to develop a targeting policy $\mathcal{P}$. First, the firm has a sample of "focal" training data available. The focal training data is obtained from an experiment in which customers were randomly assigned to a set of marketing actions A. We denote the experimental focal training data as (a_i, X_i, Y_i), $i \in \mathcal{F}$, where $\mathcal{F}$ is a sample of customers in the experiment, and a_i, X_i and Y_i indicate the randomization assignment, targeting variables and monetary outcome for customer i. In the previous example, the focal training data for the Fall Holiday catalog might be obtained from either the previous year's Fall Holiday campaign, or a small preliminary study at the start of the season.[1]

The firm also has access to one or more "source" targeting policies $\mathcal{P}_s$, $s = 1, \ldots, S$. The source targeting policies are the policies developed for different marketing campaigns. For each source campaign, the firm observes actions t_i^s recommended by each policy $\mathcal{P}_s$ for every customer $i \in \mathcal{I}$. We are agnostic to the approach and choice of estimator used to develop source policies, as long as the policies can identify a recommended action for every customer in the implementation sample $\mathcal{I}$.

The marketing actions in the source campaigns may not closely match the marketing actions in the focal problem. For example, Fall Beauty and Fall Fashion campaigns in our empirical application both include two marketing actions: *mail* and *not mail*. However, the beauty catalog is substantially different from the fashion catalog in both size and appearance. While the *no mail* actions may correspond closely, the *mail* actions are clearly different. We can also consider source and focal campaigns with more substantial differences in the action space. For example, the focal promotion may offer "$25 off", while the source promotion may offer "buy-one-get-one-free". Although the information is different, it may still indicate customer responsiveness to marketing promotions, and be relevant to choosing which action to take for the customers in the focal problem.

The requirement that the source policy can make a recommendation for every customer in the focal sample is important. In the catalog mailing setting, these restrictions will generally not introduce an obstacle when targeting existing customers. However, it may mean that if the source dataset involves existing customers, and the implementation sample is a pool of

[1] Holiday catalogs are similar in size and appearance each year. They contain the same product categories, although the specific products differ.

prospective customers, the estimation of the treatment effects from the source data cannot use measures of past purchasing, as these are not available for prospective customers.

Key Challenges

Before introducing the proposed method for transferring information from the source to the focal problems, we provide two insights to help motivate the method.

The first insight recognizes that in order to use the source policies, we need a method for identifying what information in the source policies is relevant to the focal problem. Unfortunately, because the focal training dataset is itself small, we have relatively little information to use for this identification. As we will see in the next section, this will have important implications for the performance of the resulting policy. It also has an important implication for the design of the transfer method. Because we have relatively little focal data, the method has to employ relatively few parameters to avoid overfitting.

The second insight focuses on the type of information that we transfer. We considered transferring either (a) predictions of the outcome under each marketing action, (b) predictions about the difference in outcomes between pairs of marketing actions (CATEs), or (c) policies that recommend actions for each customer. We focus on transferring either treatment effects or policies, (b) or (c), and do not recommend transferring information about mere predictions of the outcome under each marketing action (a). Targeting requires segmenting customers on treatment effects, and mere predictions about the outcome under each marketing action may include information that is not relevant to the targeting policy. If a source of variation contributes equally to the outcomes under two marketing actions, it will contribute to predictions for each action, but will not contribute to the treatment effect between these two actions (because the treatment effect measures the difference in these outcomes). For this reason, if we transfer information about predictions of the potential outcomes, we risk giving weight to information that is not relevant to the targeting decision.

Proposed Method for Transferring Information Between Problems

For ease of exposition, we present a method with three simplifying assumptions. First, we assume that there are only two possible actions in the focal and source problems, i.e. $|A| = 2$. Second, we assume a balanced focal experiment used to obtain training data, i.e. $P(a_i = 0) = P(a_i = 1), \forall i \in \mathcal{F}$. We also assume a single source problem. We discuss generalizations to these assumptions in the following subsection.

The proposed method for transferring information between marketing problems inputs the training data from the focal experiment (a_i, X_i, Y_i), $i \in \mathcal{F}$ and a source targeting policy t_i^s, and proceeds in five steps.

Step 1. Estimate the focal targeting policy on K-folds

We randomly split the sample $\mathcal{F}$ into K folds, $\mathcal{F}_k$, $k = 1, \dots, K$. For each k, we estimate a separate targeting policy $\mathcal{P}_k$ on $(K-1)$ folds and assign a recommended action on the held-out sample $\mathcal{F}_k$. In our empirical application, we assume $K = 20$ and estimate targeting policies $\mathcal{P}_k$ using the indirect Lasso method.[2] However, the choice of estimation method at this stage does not affect the remaining steps in the procedure.

The first step yields the cross-validated policy t_i^{CV} estimated on the focal dataset for all observations $i \in \mathcal{F}$. We note that for each individual customer i, the targeting variables X_i and the outcome Y_i are not used to train the prediction models for potential outcomes $\hat{Y}_{-k}(0; X)$ and $\hat{Y}_{-k}(1; X)$. This helps to avoid overfitting in Step 3.

Step 2. Calculate the expected profit for the focal, source, and inverse source targeting policies

We use the Horvitz–Thompson estimator (Horvitz and Thompson 1952) to estimate the expected profit of the cross-validated focal targeting policy, source targeting policy and the inverse source policy on the focal training data $\mathcal{F}$:

$$\pi_s = \frac{1}{|\mathcal{F}|/2} \sum_{i \in \mathcal{F}} Y_i \cdot I[t_i^s = a_i]$$

$$\pi_{\bar{s}} = \frac{1}{|\mathcal{F}|/2} \sum_{i \in \mathcal{F}} Y_i \cdot I[t_i^{\bar{s}} = a_i]$$

$$\pi_{\text{cv}} = \frac{1}{|\mathcal{F}|/2} \sum_{i \in \mathcal{F}} Y_i \cdot I[t_i^{CV} = a_i]$$

where $\bar{s}$ indicates the inverse source targeting policy $t_i^{\bar{s}} = 1 - t_i^s$. Including the inverse targeting policy recognizes that the action space in the source and focal problems might not be perfectly aligned (see later discussion).

Step 3. Estimate the combined focal targeting policy

Find $\hat{\alpha}$ to solve the following maximization problem:

$$\max_{\alpha} \frac{1}{|\mathcal{F}|/2} \sum_{i \in \mathcal{F}} Y_i \cdot I\big[w_s(\alpha) \cdot t_i^s + w_{\bar{s}}(\alpha) \cdot t_i^{\bar{s}} + w_{\text{cv}}(\alpha) \cdot t_i^{\text{cv}} \geq 0.5\big],$$

---

[2] In particular, we fit a *Lasso* model $\hat{Y}_{-k}(t, ; X)$ on $(K-1)$ folds to approximate potential outcomes $\hat{Y}_{-k}(0; X)$ and $\hat{Y}_{-k}(1; X)$, and assign recommended actions to the customers in the held-out sample using a rule $t_i^{CV} = argmax_t \, \hat{Y}_{-k}(t; X_i)$ for all $i \in \mathcal{F}_k$. The *Lasso* model includes the treatment assignment t_i, descriptive variables X_i, and the interaction terms $t_i * X_i$.

where $w_q(\alpha) = \frac{e^{\alpha\pi_q}}{e^{\alpha\pi_s}+e^{\alpha\pi_{\bar{s}}}+e^{\alpha\pi_{\text{cv}}}}$, $q \in \{s, \bar{s}, \text{cv}\}$

The objective function in this maximization problem is the expected profit under the targeting policy in which we *mail* if $w_s(\alpha) \cdot t_i^s + w_{\bar{s}}(\alpha) \cdot t_i^{\bar{s}} + w_{\text{cv}}(\alpha) \cdot t_i^{\text{cv}} \geq 0.5$, and do *not mail* otherwise. Similar to Step 2, we use the Horvitz–Thompson estimator to calculate the expected profit. The exponential terms allow us to vary the weights attributed to any number of source and focal policies, while using a single parameter α. We rely on a grid search to find the optimal value $\hat{\alpha}$.

Step 4. Estimate the focal targeting policy on the entire focal training data

We fit a *Lasso* models on the entire focal training data $\mathcal{F}$ to predict potential outcomes $\hat{Y}(0; X)$ and $\hat{Y}(1; X)$, and define a targeting policy:

$$t_i^F = argmax_t \, \hat{Y}(t; X_i) \text{ for all } i \in \mathcal{F}.$$

Compared to Step 1, we use the entire sample to design a single policy, instead of K separate policies for K cross-validation folds.

Step 5. Assign recommended actions to all customers

We use $\hat{\alpha}$ estimated in Step 3 to assign actions to customers in the implementation sample:

$$t_i^* = I\left[w_s(\hat{\alpha}) \cdot t_i^s + w_{\bar{s}}(\hat{\alpha}) \cdot t_i^{\bar{s}} + w_{\text{cv}}(\hat{\alpha}) \cdot t_i^F \geq 0.5\right] \text{ for all } i \in \mathcal{I}$$

Discussion

Our paper aims both to demonstrate that transferring information between marketing campaigns can improve targeting policies, and to investigate the boundaries of this opportunity. A systematic comparison of different methods for transferring information is beyond the scope of the paper; we see this as an important future research direction. Because different applications may require adjustments to the proposed method, we discuss several possible variants below.

The first variant is to allow estimation with multiple source policies. When there is little focal data available, the parsimony in the number of parameters is important to avoid over-fitting. We combine multiple source policies by incorporating profit-based weights in Step 3 with a single parameter α. We considered different ways to introduce regularization. For example, an alternative approach is to fix the weight for the focal dataset, and then estimate joint or separate weights for each of the source datasets. As long as the number of source datasets is not too large, this yields only a modest increase in the number of parameters, while allowing more flexibility in the weights.

Relatedly, we can consider different ways to address the ambiguity in the correspondence between the actions in the source problems and actions in the focal problem. For example, the focal problem may choose between a buy-one-get-one free promotion, and a buy-one-get-one 50% off promotion, but the source data included a $25 discount and a 20% discount. It may not be obvious which of the source promotions correspond to each of the focal promotions. This could potentially result in a negative correlation between estimates of the treatment effects from the focal and source datasets. The exponential weighting approach does not allow for negative coefficients, so we incorporate the inverse source policy into the method. Alternatively, we could fix a weight for the focal targeting policy t_i^{CV} and t_i^{F}, and then directly estimate separate weights to use for each of the source datasets.[3]

Increasing the number of parameters to flexibly incorporate multiple source campaigns and the inverse source policies can make the grid search in Step 3 computationally intensive. As an alternative to using the grid search, we could implement our transfer learning approach using a stochastic targeting policy value (Sutton et al., 1999; Silver et al., 2014). Stochastic targeting policies assign actions probabilistically for each customer, based upon the targeting variables X_i. A major advantage of the stochastic policy optimization is that we can use gradient methods to optimize $\hat{\alpha}$. However, the convergence properties of this optimization are less well-understood, and so we opt for deterministic policies in this proof-of-concept.

While we have focused on transferring targeting policies from different source problems, we could instead transfer estimated treatment effects. Targeting policies reduce the information in the treatment effect to a binary indicator identifying which policy is predicted to be optimal. By transferring policies instead of treatment effects, we may forgo relevant information. However, we also recognize that the firm cares about choosing the optimal action; and to do so it may not need to accurately predict the treatment effect.[4] Consider an example in which the estimated treatment effect for a customer is large, so that one action clearly dominates. Whether that action dominates by a large margin or a very small margin is not relevant to the design of the optimal policy. However, if the magnitude of this dominance influences the weights attributed to each dataset, this will distort the resulting policy.

An additional benefit of transferring policies is that we do not need access to the original data used to train that policy. Data from past campaigns may be inaccessible, or past policies may rely in part upon intuition or data that is not formally structured. Our method is also agnostic to the way the policy was trained. As long as we have a past policy that can be applied to the implementation data, then we can transfer information from that past policy.

[3] Another option is to check the sign of the correlation between the predictions of the source and focal treatment effects before mapping the actions in the source data to the focal actions.

[4] Targeting policies do not necessarily require estimating heterogenous treatment effects. For example, softmax value-based (Rummery and Nirajan, 1994) and policy-search methods (Williams 1992) identify optimal actions without explicitly estimating treatment effects. Transferring policies thus also increases the applicability of the proposed method.

We could use other estimators to calculate the expected profit for different targeting policies in Steps 2 and 3 (e.g., the Hajek estimator). The Horvitz-Thompson estimator is unbiased and can support estimation with unbalanced experiments by applying inverse probability weighting. The inverse probability weighting is relevant to many marketing applications, as A/B tests used to train the targeting policies often implement an unbalanced design.

Finally, marketing campaigns often send multiple versions of the promotions. For example, the firm can decide whether to *mail a $-discount*, to *mail a %-discount*, or to *not mail* to different customer segments. Our approach for transferring information between marketing problems can accommodate many actions by redefining targeting policies to compare action-specific scores in Steps 3 and 5, and adjusting the profit estimates in Steps 2 and 3. We provide an example of the estimation with three marketing actions in the Appendix.

4. Demonstrating the Value of Transferring Information

We begin by revisiting Figure 1, which we presented in the Introduction. Figure 1 demonstrates that the performance of a targeting model improves as the size of the training data increases. We produced Figure 1 using 431,159 customers in the randomly assigned treatment (*mail*) and control (*no mail*) conditions for the 2018 Fall Holiday catalog at the luxury fashion retailer (Section 2). We randomly selected 30% of the customers to use as a holdout validation sample, and then used varying subsets of the remaining customers to train targeting policies. The targeting policies selected customers to receive the catalog. We repeated the evaluation using multiple splits of the data into training and validation samples.[5]

We can use the same context to illustrate how transferring information between marketing campaigns can improve performance. In particular, we follow the procedure described in Section 3, to supplement the Fall Holiday training data using policies trained with data from the 2018 Fall Fashion catalog. The Fall Holiday and Fall Fashion catalogs featured products from different price ranges and had substantially different mailing costs due to a different page count. The business rules used to qualify customers for the two catalogs were also different, resulting in different, though overlapping, samples of customers. As a result, the targeting policy used for the Fall Fashion catalog was not optimal if directly applied to target the Fall Holiday campaign.

We report the indexed performance of the *Transfer Policy* and the *Focal Data Only* benchmark in Figure 2. The analysis in Figure 2 uses three datasets:

Source Data: the treatment and control samples from the Fall Fashion catalog.

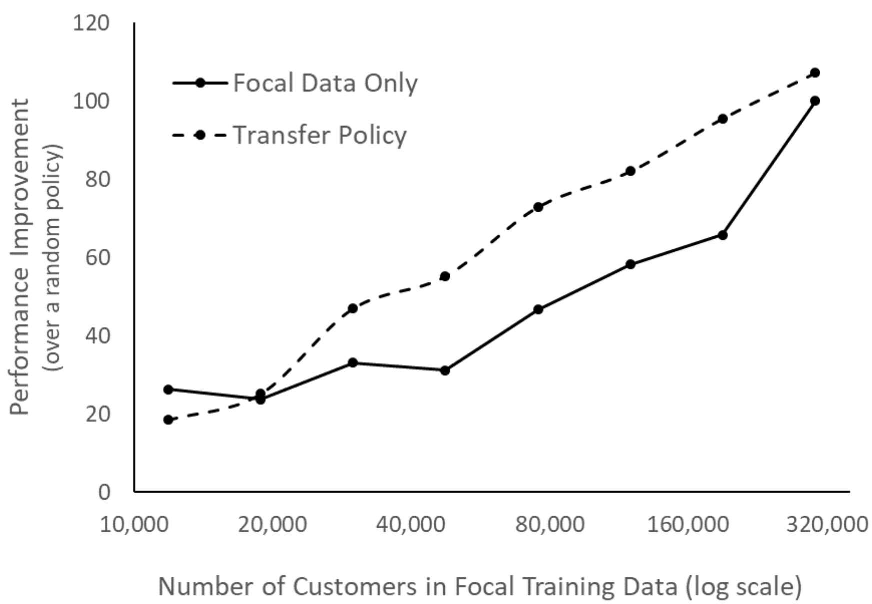
[5] See Simester et al. (2020b) for a description of how to design and evaluate a targeting policy using separate samples of experimental data.

Focal Training Data: the 70% subset of the treatment and control samples from the Fall Holiday catalog.

Validation Data: the 30% subset of the treatment and control samples from the Fall Holiday catalog.

The *Transfer Polic*y is a policy in which we transferred information from the Fall Fashion catalog to improve targeting for the Fall Holiday catalog; the *Focal Data Only* benchmark is trained using only the focal data. To investigate where the *Transfer Policy* outperforms the *Focal Data Only* benchmark, we randomly select different-sized subsamples of the focal training data and use the x-axis in Figure 2 to denote the size of the focal training dataset. The size of the source data and the validation data are both fixed.

Figure 2. Indexed Performances of the *Focal Data Only* and *Transfer Policy*



The figure reports the performance improvement of a *Focal Data Only* policy trained using only the focal data, and a *Transfer Policy* that transfer information from a separate source campaign. The performance improvements are calculated compared to a random policy. Different sized focal data training samples were randomly selected, and the x-axis indicates the sample size of the training data (using a log scale). The outcome measures are indexed by setting the performance of the *Focal Data Only* policy to 100 when the focal data contains 300,000 observations.

The performance of both policies is compared to the random policy, which assigns *mail* and *no mail* decisions to each customer with equal probability. We indexed the performance by setting the performance improvement of the *Focal Data Only* policy to 100 when the sample size of the focal training data is largest (300,000 observations).

Figure 2 reveals that the *Transfer Policy*, where we supplemented the focal training data with information from a different campaign, consistently outperforms policies trained using only the focal training data. The improvement is particularly notable as the customers in the focal training data and validation data are equivalent (due to randomization), and the marketing actions in the focal training data and validation data are identical (by construction). As a result, the *Focal Data Only* policy is a relatively strong benchmark. Still, Figure 2 demonstrates that data from a different experiment involving mailing a different catalog to customers selected using different business rules, can help to improve the targeting policy.

The improvements are important. As we discussed in the Introduction, the *Transfer Policy* with 50,000 customers yields the same performance as the *Focal Data Only* policy with almost 120,000 customers. This represents a 58% reduction in the required sample size.

The performance improvement in Figure 2 depends upon how much focal training data is available. We illustrate this more clearly in Figure 3, by reporting how the difference in the performance of the *Transfer Policy* and *Focal Data Only* benchmark varies with the size of the focal training data.

Figure 3. Performance Improvement and the Size of the Focal Training Dataset

The figure reports the difference in the performance of the *Transfer Policy* and the *Focal Data Only* policy when varying the number of observations in the focal training data. The performance of each policy is illustrated separately in Figure 2. The x-axis indicates the sample size of the focal training data (using a log scale). The outcome measures are indexed by setting the performance of the *Focal Data Only* policy to 100 when the focal training data contains 300,000 observations.

The performance improvement exhibits an inverse-U shaped pattern as the sample size of the focal training data varies. When the sample size of the focal dataset is too small, there is not enough information to identify what variation in the source data is relevant to the focal problem. For this reason, even though the source data contains relevant information, we are not able to isolate this information and use it to improve the targeting policy. At the other extreme, when there is a lot of training data, the source data provides little incremental information that is relevant to the focal problem, beyond the information contained in the focal data itself.

The inverse U-shape relationship holds even if we consider an ideal (hypothetical) setting where the source data is identical to the focal data. If the amount of focal training data is too small, we cannot distinguish whether the relevance of the information in the source data is high or low. As the amount of focal training data becomes smaller, the weights attributed to the source data and focal data (in Step 3) become random, which undermines the quality of the combined policy. At the other extreme, if the focal training data is very large, the incremental value of the source data (even ideal source data) diminishes as the room for improvement becomes smaller.

The implication is that transferring information from source datasets is most likely to improve the performance of a targeting policy when the focal data is small – but not too small. We need enough information about the focal problem to know what source information to transfer. However, the information in the focal dataset must be at least somewhat incomplete, otherwise the source information provides no incremental value.

5. Pre-Screening Potential Source Policies

Recall that our proposed approach for transferring information estimates the value of the source targeting policy π_s and the value of the inverse source policy $\pi_{\bar{s}}$ when these policies are directly applied to the focal problem, without adjustments. Targeting policies that yield a higher profit on the focal data sample receive a higher weight in the combined policy. Intuitively, $V_s = \max(\pi_s, \pi_{\bar{s}})$ characterizes the amount of information about the focal problem that is contained in the source targeting policy. We refer to V_s as the *Direct Transfer Value* and propose using this measure to prescreen which past campaigns are attractive candidates for use as source data.[6]

[6] This measure was motivated in part by a method commonly used in the direct mail industry for choosing mailing lists when prospecting for new customers. The characteristics of candidate mailing lists vary on many dimensions, and it is not obvious which characteristic is most informative about fit with the firm. Instead of choosing mailing lists based upon the characteristics of the mailing lists, many firms request a small random sample of customers from the mailing list. If the overlap between this sample and the firm's existing customers is high (i.e. over 5% of the random sample are already customers), this indicates that the fit of the mailing list is high.

large. As we have discussed, small focal training data does not identify the informative signal in the source policies, and the resulting *Transfer Policy* is no better than random. Alternatively, transferring information with large focal training data provides little performance improvement, because the source policies contain little incremental information over the information already in the focal data.

For a moderately large focal training data, we observe performance gains due to transferring information when *Transfer Policies* are less noisy. The source policy without induced random noise yields the largest difference between the *Focal Data Only* and *Transfer Policy* for any considered sample size. On the other side, transferring a fully random source policy yields a policy which is worse than *Focal Data Only*. A random source policy adds noise and no informative signal to the focal training data and undermines performance (due to the finite samples).

Non-informative source policies present a challenge for transferring information with small focal experiments. The focal training data may be insufficient to distinguish informative and non-informative signals. The *Direct Transfer Value*, V_s, can help to pre-screen which source campaigns are more likely to provide valuable information about the focal problem. This is illustrated by the performance of the perturbed policies in Figure 4. With no random component, the *Direct Transfer Value* $V_s = 2.0$, indicating that direct transfer outperforms the benchmark random policy by 2%. As the percentage of random recommendations in the source policy increases to 25%, 50%, 75% and 100%, the value of V_s decreases to 1.4, 0.8, 0.2 and -0.3 (respectively).[8]

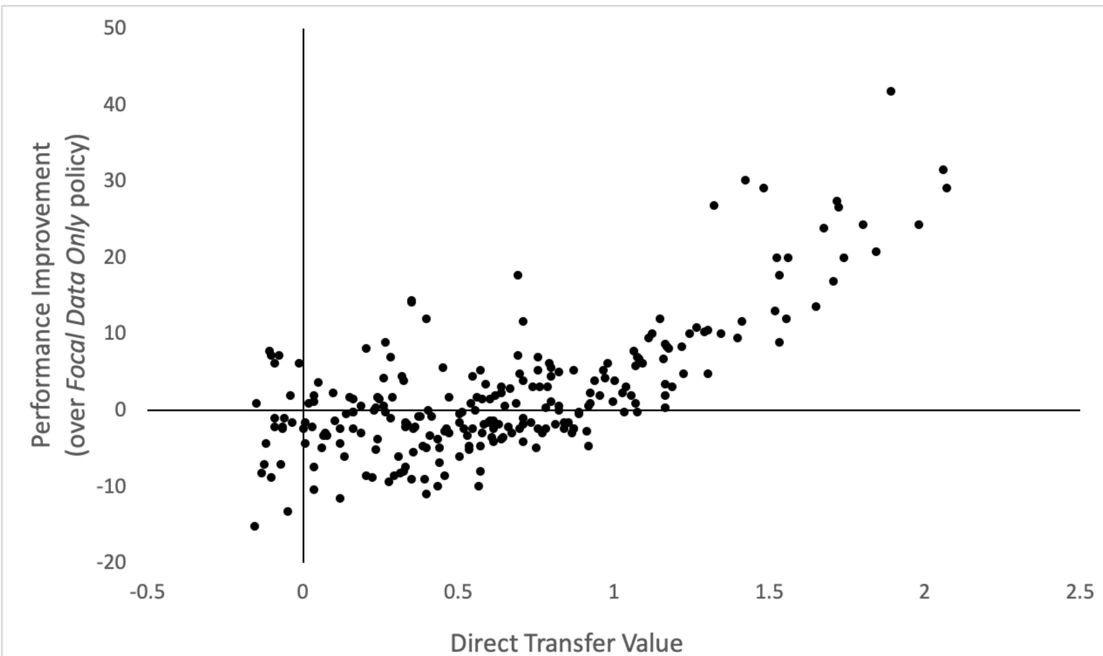
In Figure 5, we provide additional evidence that the *Direct Transfer Value* can help to pre-screen source policies. Recall that our data includes many direct mail campaigns conducted by the luxury fashion retailer in 2017-2019. We use the 2018 Fall Holiday catalog as the focal campaign. As potential source datasets, we evaluate 242 campaigns that concluded prior to the in-home date for this focal campaign. For every source campaign, we estimate a *Direct Transfer Value* and the performance improvement of the *Transfer Policy* over the *Focal Data Only* policy. In Figure 5, we plot the performance improvement when separately transferring each of these source campaigns to the focal campaign. Each point in the figure represents one of the 242 source campaigns. The x-axis describes the *Direct Transfer Value* for each campaign, while the y-axis reports the performance improvement when transferring information to the focal problem (compared to just using the focal data alone). The focal training data includes 190,000 customers, which represents a moderately large focal training data. As a robustness check, we repeated the analysis using 80,000 customers in the focal training data. The findings are consistent.

[8] For confidentiality reasons, we scale the estimated profit by the performance of the random policy.

The pair-wise correlation between the *Direct Transfer Value* and the performance improvement is 0.71; source policies with larger *Direct Transfer Values* tend to result in larger performance gains. In our application, all source policies with a *Direct Transfer Value* exceeding 1.08 yield a positive performance improvement, and provide good candidates for transferring information.[9]

Figure 5 also confirms our earlier observation that transferring information from non-informative source campaigns can worsen performance of the targeting policies. *Transfer Policies* with low *Direct Transfer Values* often perform worse than the *Focal Data Only* policy. Non-informative source campaigns add noise to the estimation approach in the focal campaign, so the performance deteriorates with finite samples.

Figure 5. Value of the Direct Transfer and Robustness to Randomized Policies



The figure compares the direct transfer value and the performance improvement of the *Transfer Policy* over the *Focal Data Only* policy for different source datasets. The focal campaign is the 2018 Fall Holiday catalog and the focal training data includes 190,000 customers. The *Transfer Policies* are trained using source policies trained for different marketing campaigns. The performance improvement is indexed by setting the performance of the *Focal Data Only* policy to 100 when the focal training data contains 300,000 observations. The *Direct Transfer Value* indicates the percent improvement of direct transfer over the random benchmark. Each point represents a single source policy, and the picture includes 242 points.

Table 1 provides examples of several source campaigns with large *Direct Transfer Values*. The range of campaigns is notable. Recall that the focal campaign in our analysis is the 2018 Fall Holiday catalog. As expected, the 2017 Fall Holiday catalog is among the most informative

[9] In our setting, the 1.08% improvement over the random benchmark corresponds to over $1 in incremental profit per customer.

source campaigns. Retailers would typically use this campaign as a training data for the new targeting policy in fall 2018. Other informative campaigns include the Fashion catalogs, Father's Day catalogs, together with postcards offering invitations to special events, promotions designed to reactivate lapsed customers, and loyalty card promotions. These campaigns communicate different marketing messages, have different formats (postcard or catalogs), and address different customer strata.

Table 1. Examples of Informative Source Campaigns

Description	Format	Season	Direct Transfer Value	Performance Improvement (Figure 5)
Holiday Catalog	Catalog	Fall 2017	1.89	$41.71
Fashion Catalog	Catalog	Spring 2017	1.41	$11.60
		Fall 2017	1.98	$24.31
		Spring 2018	1.80	$24.31
Father's Day Catalog	Catalog	Spring 2017	2.07	$29.01
		Spring 2018	1.32	$26.80
Invitation to the Loyalty Program	Postcard	Summer 2017	1.34	$9.94
Reactivate Lapsed Customers	Postcard	Spring 2018	1.40	$9.39
Final Sale and Friends & Family Events	Postcard	Spring 2017	1.52	$19.89
		Summer 2018	1.42	$30.11
Valentine's Day Event	Postcard	Winter 2017	1.72	$27.35

The table describes the characteristics of several source campaigns with large *Direct Transfer Values* when the focal campaign is the 2018 Fall Holiday catalog. The focal training data includes 190,000 customers.

We further explore how characteristics of the source policies relate to the performance improvements in Table 2. We consider the 242 source marketing campaigns that concluded prior to the focal problem (Figure 5) and provide a regression analysis of the performance improvement on the observed characteristics of the marketing campaigns. The characteristics include the time difference between the source and focal campaign in-home dates, the number of customers in the source campaigns, and the format of the direct mail campaign (catalog or postcard).

Table 2. Performance Improvement for Different Source Marketing Campaigns

	DV: **Performance Improvement** (over *Focal Data Only* policy)		
	(1)	(2)	(3)
Intercept	-5.709**	-5.879†	-11.868**
	(0.636)	(3.340)	(2.393)
Direct Transfer Value	12.228**		12.059**
	(0.787)		(0.780)
Time Between Source and Focal Campaigns (in months)		-0.041	-0.033
		(0.104)	(0.074)
Number of Customers in Source Campaign (log)		0.792**	0.635**
		(0.294)	(0.208)
Indicator for Catalog		2.237	0.713
		(1.611)	(1.143)
R^2	0.502	0.039	0.522

The table reports the relationship between the characteristics of the source targeting policies and the performance improvement attributable to information transfer. The focal campaign is the 2018 Fall Holiday catalog and the focal training data includes 190,000 customers. The analysis includes 242 source campaigns concluded prior to the focal problem. Standard errors are in parentheses. **Indicates significantly different from zero (p < 0.01). *Indicates significantly different from zero (p < 0.05). †Indicates significantly different from zero (p < 0.10).

Table 2 confirms that the *Direct Transfer Value* provides a strong signal of which source campaigns yield the largest profit improvements. A linear model with the intercept and the *Direct Transfer Value* explains half of the variance in the performance improvement ($R^2 = 0.502$; Column 1).

In Column 2, we observe that the performance improvement tends to be larger for the source campaigns that contain more customers. The performance improvement from the campaigns that are more recent and that have the same format as the focal campaign also tend to be larger, but these relationships are not statistically significant. Collectively, the descriptive characteristics of the source campaigns provide a weaker signal than the *Direct Transfer Value* about which source policies will yield the largest performance improvement (compare R^2 in Column 3 and Column 1). We conclude that the *Direct Transfer Value* can serve as a useful indicator to pre-screen which source campaigns to focus on when training a targeting policy for a new campaign.

6. Replicating the Findings at Two Different Firms

To demonstrate the robustness of our results, we replicate the primary findings at the luxury fashion retailer using data from two different firms: a membership wholesale club and a large financial services firm. There are many differences in the targeting problems faced by these three firms. The differences extend beyond just their industries, to include differences in the stage of the customer life cycle, the value propositions in the marketing actions, the marketing channels, the size of the treatment effects, the size of the action space, and the unit of analysis. We summarize these differences in Table 3 (additional details are provided in the Appendix).

Table 3. Differences Across the Three Firms

	Luxury Fashion Retailer	Membership Wholesale Club	Financial Services Firm
Industry	Luxury fashion	Discount groceries and hard goods	Financial services
Campaign Goal	Sell next season's products to existing customers	Attract new customers	Information about products and services
Target market	Existing customers	Prospective customers	Existing customers
Unit of Analysis	Customer	Zip Code	Customer
Value Proposition	Information about new products	Discounted memberships	Information about additional services
Marketing Channel	Direct mail	Direct mail	Email
Treatment Effects	Large	Medium	Small
Action Space	Mail, No Mail	Mail Promotion A, Mail Promotion B, No Mail	Mail, No Mail

The replications follow the analysis presented in Section 4. For each firm we identify a single focal campaign with experimental variation in treatment and control assignment. We then iteratively randomly divide the focal data into focal training and validation (holdout) data, train *Focal Data Only* policies using different sized subsamples of the focal training data, and evaluate the resulting policies on the holdout data. We then identify a separate campaign to use as a source dataset, and transfer information from the source data to the focal problem, using the method described in Section 3. We refer to the policies that incorporate the transferred information as *Transfer Policies*.

In the membership wholesale club analysis, the focal and source campaigns both targeted prospective customers with postcard promotions. Both campaigns included three marketing actions (two types of promotional offers and a no mail control). However, there were also important differences between the focal and source campaigns. In the focal campaign, the promotions were sent by postcard, while in the source campaign, the promotions were highlighted on the front and back covers of a 48-page book of discounted coupons. Moreover, the geographic coverage was different. Customers in the source dataset were concentrated in two geographic regions, while customers in the focal dataset were spread across 100 regions (just 2% of the regions overlapped).

In the financial services study, the focal and source campaigns informed existing customers about the same new products and services. The two campaigns were conducted in different marketing channels: the source campaign was conducted using email, while the focal campaign was conducted using direct mail. The customer bases were also different, with the email campaign including substantially more customers than the direct mail campaign.

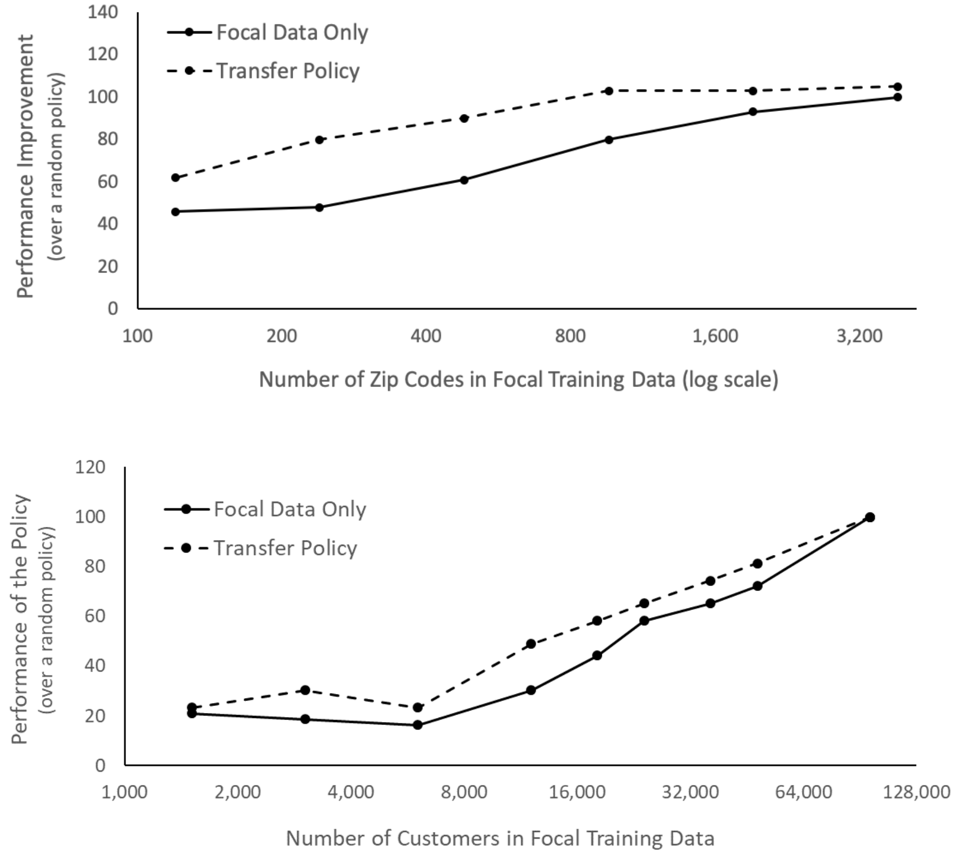
Because the experiments at the wholesale membership club focus on prospective customers, who have no purchasing histories, the company only had access to aggregate demographic measures to train a targeting model. We aggregated observations and the outcome measures to the zip code level, and the analysis was conducted at this level. At the luxury goods retailer and the financial services firm, the analysis was all conducted at the customer level.

Figure 6 compares the performance of the targeting policies with and without information transfer. The findings in these replications mimic the pattern in Section 4. Transferring data from a separate source problem improves the performance of targeting policies at both firms. The performance improvement is particularly large when the size of the focal data is moderately large. As we discussed, transferring information from a separate source problem requires enough information about the focal problem to know which information in the source dataset is relevant to the focal problem. However, if the size of the focal training data is very large, there is less incremental value provided by the information in the source dataset.

Figure 6. Replications at Two Additional Firms

(a) Wholesale Membership Club

(b) Financial Services Firm



The figure reports the performance of a *Focal Data Only* policy trained using only the focal data, and a *Transfer Policy* that transfer information from a separate source campaign. The focal campaign is the 2018 Fall Holiday catalog. The performance improvements are calculated compared to a random policy. The x-axis indicates the sample size of the training data (different sized samples were randomly selected). The outcome measures are indexed by setting the performance of the *Focal Data Only* policy to 100 when the focal data contains the entire available sample.

The replication is particularly reassuring given the many differences in the targeting problems faced by the three firms (Table 2). It suggests that the benefits of the proposed approach generalize to different marketing settings.

7. Conclusion

Firms face a challenge when designing experiments to train targeting policies: large experiments provide more training data and yield better targeting performance, but large experiments are often costly or infeasible. We proposed an approach to train targeting policies by supplementing small experimental training data with information from prior marketing campaigns. Our approach combines information across marketing campaigns to improve the performance of targeting policies. We apply the proposed approach to data from a luxury fashion retailer and demonstrate the robustness of the effect by replicating the findings at a financial services firm and a membership wholesale club.

Transferring information between marketing campaigns is particularly valuable when the focal experimental data is neither too small nor too large. If there is too little focal data, the methods cannot identify what information in the source targeting policies is relevant to the focal problem. The benefits of transferring information are also eroded if the sample size of the focal training data is large, so that there is less potential for incremental improvement.

The value of information transfer also depends upon the relevance of the source marketing campaign(s) to the focal targeting problem. Incorporating non-informative marketing campaigns into the estimation approach with small focal training data can worsen the performance of the targeting models. We propose a profit-based measure to pre-screen potential source campaigns.

8. References

Ascarza, Eva. "Retention futility: Targeting high-risk customers might be ineffective." *Journal of Marketing Research* 55.1 (2018): 80-98.

Athey, Susan, and Guido Imbens. "Recursive partitioning for heterogeneous causal effects." *Proceedings of the National Academy of Sciences* 113.27 (2016): 7353-7360.

Bubeck, Sébastien, and Nicolo Cesa-Bianchi. "Regret analysis of stochastic and nonstochastic multi-armed bandit problems." *arXiv preprint arXiv:1204.5721* (2012).

Burnap, Alex, John R. Hauser, and Artem Timoshenko. "Design and evaluation of product aesthetics: a human-machine hybrid approach." *Available at SSRN 3421771*(2019).

Caselles-Dupré, Hugo, Florian Lesaint, and Jimena Royo-Letelier. "Word2vec applied to recommendation: Hyperparameters matter." *Proceedings of the 12th ACM Conference on Recommender Systems*. 2018.

Chernozhukov, Victor, et al. *Generic machine learning inference on heterogenous treatment effects in randomized experiments*. No. w24678. National Bureau of Economic Research, 2018.

Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018).

Dietterich, Thomas G. "Ensemble methods in machine learning." *International workshop on multiple classifier systems*. Springer, Berlin, Heidelberg, 2000.

Donahue, Jeff, et al. "Decaf: A deep convolutional activation feature for generic visual recognition." *International conference on machine learning*. 2014.

Dubé, Jean-Pierre, and Sanjog Misra. *Scalable price targeting*. No. w23775. National Bureau of Economic Research, 2017.

Gittins, John, Kevin Glazebrook, and Richard Weber. *Multi-armed bandit allocation indices*. John Wiley & Sons, 2011.

Grimmer, Justin, Solomon Messing, and Sean J. Westwood. "Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods." *Political Analysis* 25.4 (2017): 413-434.

Guelman, Leo, Montserrat Guillén, and Ana M. Pérez-Marín. "Random forests for uplift modeling: an insurance customer retention case." *International Conference on Modeling and Simulation in Engineering, Economics and Management*. Springer, Berlin, Heidelberg, 2012.

Hájek, J. "Comment on An essay on the logical foundations of survey sampling by Basu, D." *Foundations of Statistical Inference* 236 (1971).

Hitsch, Guenter J. and Misra, Sanjog, Heterogeneous Treatment Effects and Optimal Targeting Policy Evaluation (January 28, 2018). Available at SSRN: https://ssrn.com/abstract=3111957 or http://dx.doi.org/10.2139/ssrn.3111957

Lewis, Randall A., and Justin M. Rao. "The unfavorable economics of measuring the returns to advertising." *The Quarterly Journal of Economics* 130.4 (2015): 1941-1973.

Li, Jimmy Q., et al. "The value of field experiments." *Management Science* 61.7 (2015): 1722-1740.

Li, Lihong, et al. "A contextual-bandit approach to personalized news article recommendation." *Proceedings of the 19th international conference on World wide web*. 2010.

Liu, Liu, Daria Dzyabura, and Natalie Mizik. "Visual listening in: Extracting brand image portrayed on social media." *Marketing Science* 39.4 (2020): 669-686.

Mikolov, Tomas, Quoc V. Le, and Ilya Sutskever. "Exploiting similarities among languages for machine translation." *arXiv preprint arXiv:1309.4168* (2013).

Ng, Hong-Wei, et al. "Deep learning for emotion recognition on small datasets using transfer learning." *Proceedings of the 2015 ACM on international conference on multimodal interaction*. 2015.

Nichols, Austin. "Causal inference with observational data." *The Stata Journal* 7.4 (2007): 507-541.

Oquab, Maxime, et al. "Learning and transferring mid-level image representations using convolutional neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.

Pan, Sinno Jialin, and Qiang Yang. "A survey on transfer learning." *IEEE Transactions on knowledge and data engineering* 22.10 (2009): 1345-1359.

Peysakhovich, Alexander, and Akos Lada. "Combining observational and experimental data to find heterogeneous treatment effects." *arXiv preprint arXiv:1611.02385* (2016).

Rafieian, O. (2019), “Optimizing User Engagement through Adaptive Ad Sequencing,” working paper, University of Washington.

Rafieian, O. and H. Yoganarasimhan (2018), “Targeting and Privacy in Mobile Advertising,” working paper, University of Washington.

Rummery, G. A. and Niranjan, M. On-line Q-learning using connectionist systems. Technical Report CUED/F-INFENG/TR 166, Cambridge University Engineering Department, 1994.

Schwartz, Eric M., Eric T. Bradlow, and Peter S. Fader. "Customer acquisition via display advertising using multi-armed bandit experiments." *Marketing Science* 36.4 (2017): 500-522.

Siirndal, C., Swensson,B., and Wretman, J. (1992), Model Assisted Survey Sampling, Springer Verlag, New York.

Silver, David, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. "Deterministic policy gradient algorithms." 2014.

Simester, Duncan, Artem Timoshenko and Spyros Zoumpoulis (2020a), “Targeting Prospective Customers: Robustness of Machine Learning Methods to Typical Data Challenges,” *Management Science*, 66(6), 2291-2799.

Simester, Duncan, Artem Timoshenko, and Spyros I. Zoumpoulis (2020b). "Efficiently Evaluating Targeting Policies: Improving on Champion vs. Challenger Experiments." *Management Science*, 66(8), 3295-3798.

Sutton, R. S., McAllester, D. A., Singh, S. P., and Man- sour, Y. (1999). Policy gradient methods for reinforce- ment learning with function approximation. In *Neural Information Processing Systems 12*, pages 1057–1063.

Wang, Xiaosong, et al. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Williams, Ronald J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3):229–256, 1992.

Yoganarasimhan, Hema, Ebrahim Barzegary, and Abhishek Pani. "Design and Evaluation of Personalized Free Trials." *arXiv preprint arXiv:2006.13420* (2020).

Zhou, Zhi-Hua. *Ensemble methods: foundations and algorithms*. CRC press, 2012.

Zou, Will Y., et al. "Bilingual word embeddings for phrase-based machine translation." *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. 2013.

Appendix A. Replication at the Membership Wholesale Club

The membership wholesale club retailer in our replication analysis is a large chain of US wholesale clubs. Customers are required to have memberships in order to shop at this retailer's stores. We use data from two experiments, both of which were designed to encourage new customers to sign up for memberships.

In both experiments, customers were randomly assigned to receive one of three marketing actions:

1. Promotion 1: $25 discounted membership
2. Promotion 2: 120-day free-trial membership
3. Control: no intervention

The source experiment used in our replication analysis was conducted in Spring 2015. The promotions were highlighted on the front and back cover of a 48-page catalog of discounted coupons. The treatments were repeated twice, approximately six weeks apart (customers in the promotion conditions received the same promotion offer twice). The customers were all located in zip codes clustered around two of the regions in which the retailer had stores. Additional details of this experiment are described in Simester et al. (2020a).

The focal experiment was conducted approximately 12-months later, in Spring 2016. In this experiment, the promotions were sent to prospective customers using a postcard, which was printed on both sides, and highlighted the offer on each side. Customers in the focal experiment were located in a much larger set of regions, with just 2% of the households in the focal experiment located in the same geographic area as the source experiment. The treatments were again repeated twice, although in this case the second wave of mailings was just three weeks after the first wave.

The source data experiment included a total of 412 zip codes, while the focal data experiment included 2,558 zip codes. The 2,558 zip codes and three treatments in the focal data yielded 7,674 data points. These observations were randomly grouped into two sub-samples: 3,837 in the *Focal Training Data* and 3,837 in the *Validation Data*.

In both experiments, the outcome variable measured the average profit earned from each customer (we assumed a mailing cost of $0.45). This customer-level variable was aggregated up to the 5-digit zip code level, and all of the analysis was conducted at the zip code level. For each experiment, the initial targeting policies were estimated using *Lasso*. The source and focal models included the same set of ten demographic variables.

Appendix B. Replication at the Financial Services Firm

Our replication analysis uses data from a large US financial institution. The data contains anonymized, account-level information about customers' financial holdings.

In both experiments, existing customers were randomly assigned to receive one of two marketing actions:

1. Treatment: information about financial products and services
2. Control: no intervention

The source experiment was conducted in April 2017. In this experiment, the information was provided to customers in the treatment condition through an email. The sample size was 211,642 households.

The focal experiment was conducted in the same month and year as the source experiment (April 2017), and included information about the same financial products and services. In this experiment, the information was provided to customers in the treatment condition using a letter sent by US mail. The customers in the source and focal experiments did not overlap. The sample size in the focal experiment was 137,837, which was randomly grouped into two sub-samples: 96,486 (*Focal Training Data*) and 41,351 (*Validation Data*).

In both experiments, the outcome measure was a continuous customer-level measure of profit. The initial targeting policies were estimated using gradient boosted trees (*LightGBM*). The source and focal models included the same set of variables, describing historical changes in each customer's financial holdings (analogous to RFM variables in retail purchasing data).

Appendix C. Example with Three Marketing Actions

In our replication study with the membership wholesale club, we modify the estimation approach proposed in Section 3 to accommodate three marketing actions in the source and focal campaigns. The primary adjustment is the definition of the targeting policies in Steps 3 and 5. We assume that marketing actions are numerated or $A = \{0,1,2\}$.

Step 1. Estimate the focal targeting policy on K-folds

We randomly split the sample $\mathcal{F}$ into K folds, $\mathcal{F}_k$, $k = 1, \dots, K$. For each k, we estimate a separate targeting policy $\mathcal{P}_k$ on $(K-1)$ folds and assign a recommended action on the held-out sample $\mathcal{F}_k$.

Step 2. Calculate the expected profit for the focal, source, and inverse source targeting policies

We use the Horvitz–Thompson estimator (Horvitz and Thompson 1952) to estimate the expected profit of the cross-validated focal targeting policy and the permutations of the source policy on the focal training data $\mathcal{F}$:

$$\pi_{s^p} = \frac{1}{|\mathcal{F}|/3} \sum_{i \in \mathcal{F}} Y_i \cdot I[t_i^{s^p} = a_i]\,, p = 1, \dots, 6$$

$$\pi_{\text{cv}} = \frac{1}{|\mathcal{F}|/3} \sum_{i \in \mathcal{F}} Y_i \cdot I[t_i^{CV} = a_i]$$

where s^p indicates the permutations of the source targeting policy t_i^s. Permutations of the source targeting policy include all possible assignments between the actions in the source and focal marketing campaigns. For example, in our case of three marketing actions in the source and focal campaigns there are 6 possible permutations.

Step 3. Estimate the combined focal targeting policy

Find $\hat{\alpha}$ to solve the following maximization problem:

$$\max_{\alpha} \frac{1}{|\mathcal{F}|/3} \sum_{i \in \mathcal{F}} Y_i \cdot I\left[t_i^{Step\ 3}(\alpha) = a_i\right],$$

where

$$t_i^{Step\ 3}(\alpha) = argmax \begin{pmatrix} \sum_p w_{s^p}(\alpha) \cdot I[t_i^{s^p} = 0] + w_{\text{cv}}(\alpha) \cdot I[t_i^{CV} = 0] \\ \sum_p w_{s^p}(\alpha) \cdot I[t_i^{s^p} = 1] + w_{\text{cv}}(\alpha) \cdot I[t_i^{CV} = 1] \\ \sum_p w_{s^p}(\alpha) \cdot I[t_i^{s^p} = 2] + w_{\text{cv}}(\alpha) \cdot I[t_i^{CV} = 2] \end{pmatrix}$$

$$w_q(\alpha) = \frac{e^{\alpha \pi_q}}{\sum_p e^{\alpha \pi_{s^p}} + e^{\alpha \pi_{\text{cv}}}}, \qquad q \in \{s^1, \dots, s^6, \text{cv}\}$$

The objective function in this maximization problem is the expected profit under the targeting policy which takes the best action amongst 3 possible ones.

Step 4. Estimate the focal targeting policy on the entire focal training data

We fit a *Lasso* models on the entire focal training data $\mathcal{F}$ to predict potential outcomes $\hat{Y}(0; X)$, $\hat{Y}(1; X)$, and $\hat{Y}(2; X)$ and define a targeting policy:

$t_i^F = argmax_t\, \hat{Y}(t; X_i)$ for all $i \in \mathcal{F}$.

Compared to Step 1, we use the entire sample to design a single policy, instead of K separate policies for K cross-validation folds.

Step 5. Assign recommended actions to all customers

We use $\hat{\alpha}$ estimated in Step 3 to assign actions to customers in the implementation sample:

$$t_i^*(\hat{\alpha}) = argmax \begin{pmatrix} \sum_p w_{s^p}(\hat{\alpha}) \cdot I[t_i^{s^p} = 0] + w_{\mathrm{cv}}(\hat{\alpha}) \cdot I[t_i^{CV} = 0] \\ \sum_p w_{s^p}(\hat{\alpha}) \cdot I[t_i^{s^p} = 1] + w_{\mathrm{cv}}(\hat{\alpha}) \cdot I[t_i^{CV} = 1] \\ \sum_p w_{s^p}(\hat{\alpha}) \cdot I[t_i^{s^p} = 2] + w_{\mathrm{cv}}(\hat{\alpha}) \cdot I[t_i^{CV} = 2] \end{pmatrix}$$

$$\text{for all } i \in \mathcal{I}$$