Lookalike Targeting on Others' Journeys: Brand Versus Performance Marketing

K. Sudhir* Seung Yoon Lee* Subroto Roy†

June 2022

Abstract

Lookalike targeting is a widely used model-based ad targeting approach that uses a seed database of individuals to identify matching "lookalikes" for targeted customer acquisition. An advertiser has to make two key choices: (1) who to seed on and (2) seed-match rank range. First, we find that seeding on others' journey stage can be effective in new customer acquisition; despite the cold start nature of customer acquisition using Lookalike audiences, third parties can indeed identify factors unobserved to the advertiser that move individuals along the journey and can be correlated with the lookalikes. Further, while journey-based seeding adds no incremental value for brand marketing (click-through), seeding on more downstream stages improves performance marketing (donation) outcomes. Second, we evaluate audience expansion strategies by lowering match ranks between the seed and lookalikes to increase acquisition reach. The drop in effectiveness with lower match rank range is much greater for performance marketing than for brand marketing. Performance marketers can alleviate the problem by making the ad targeting explicit, and thus increase perceived relevance; however, it has no incremental impact for higher match lookalikes. Increasing perceived targeting relevance makes acquisition cost comparable for both high and low match ranks.

Keywords: digital advertising, targeting, algorithmic targeting, lookalike targeting, nonprofit marketing

*Yale School of Management; k.sudhir@yale.edu, seungyoon.lee@yale.edu

†University of New Haven; roy@newhaven.edu

The authors thank Vineet Kumar, Soheil Ghili, Aniko Oery, Jiwoong Shin, Kosuke Uetake and participants at the Yale SOM Lunch Workshop and the IIM Bangalore Behavioral Sciences Lab Webinar, IIM Jammu for comments and helpful discussions. We also thank Anja Lambrecht, Stephan Seiler and Song Yao for constructive feedback on an earlier draft.

INTRODUCTION

Lookalike advertising is a targeting approach in digital advertising where the advertiser uses a "seed database" of customers with desired behaviors to algorithmically identify "matching lookalikes" in a much larger third party database for targeting. The technique, originally introduced by Facebook in 2013, helps advertisers target and acquire customers using the depth and breadth of third party data available to Facebook, by finding similarities with customers that exhibit "desirable" behaviors (e.g., browse, visit, purchase, donate). The assumption behind the targeting technique is that greater similarity (correlation) in behaviors and descriptors (e.g., demographics/psychographics) with a focal firm's desirable seed individuals leads to greater advertising responsiveness and more efficient acquisition. Several large data-driven advertising platforms such as Google, Twitter, LinkedIn, Outbrain and Taboola also offer Lookalike ad targeting services.[1]

Lookalike targeting is considered most useful to expand customer acquisition to new "cold" audiences; it thus differs from many digital targeting methods that improve efficiency by targeting "warm" customers who are either on the journey to purchase or have already purchased. For example, keyword search (e.g., Rutz and Bucklin 2011) uses information on a customer's needs captured in search terms; retargeting (e.g., Lambrecht and Tucker 2013) uses knowledge of products browsed, and contextual targeting (e.g., Ghose, Li, and Liu 2019) uses the journey context of the focal individual to target. A focal individual's past purchasing/browsing behaviors (e.g., Rossi, McCulloch, and Allenby 1996, Pancras and Sudhir 2007, Rafieian and Yoganarasimhan 2020) is also well-known to improve targeting effectiveness. But targeting only on warm customers is often insufficient to meet a firm's customer acquisition volume goals. Lookalike targeting promises to serve a complementary role to these methods by efficiently expanding customer acquisition to cold prospects who are similar to the warm prospects already on the journey by "kickstarting" their interest.

[1] Google calls its service "Similar Audiences." Google claims Lookalike Targeting improves performance relative to standard digital ad networks based on firm-specified targeting by over 41%. (https://www.thinkwithgoogle.com/future-of-marketing/emerging-technology/similar-audiences)

While it is reasonable to expect improvements in targeting based on one's own journey moving down the purchase funnel, whether others' journey stage can serve as effective seeds for Lookalike targeting is not a priori obvious. This is because the profiled lookalikes based on journey stage seeding have never themselves embarked on the journey with the focal firm. On the one hand, journey-based seeding may enable advertisers to identify lookalikes that are already interested in the category or engage with similar brands and competitors. But, to the extent that movement along the journey to purchase is based on brand-specific and contextual needs unrelated to permanent (and observed to third party) characteristics of the seed that are correlated with the lookalike, journey-based lookalike seeding would unlikely add incremental value for customer acquisition.

This paper offers the first empirical investigation of Lookalike targeting and provides guidance for its practical use. First, we evaluate the fundamental premise of Lookalike targeting: can seeding lookalikes based on journey stage of *others* indeed improve targeting effectiveness? And if effective, what journey stage should an advertiser use? For an advertiser seeking an upstream behavior (click-through to a brand site for brand building), would there be incremental value in seeding on behaviors further down the journey (e.g., purchase, donation), or is seeding on the upstream clickthrough behavior sufficient? Similarly, for an advertiser seeking downstream performance outcomes, would there be an incremental value in seeding further down the journey on loyalty (e.g, repeat purchase, lifetime value, WOM), versus seeding on single past purchase event? On the one hand, due to favorable selection inherent in moving down the journey, seeds from further down the customer journey can improve targeting, regardless of brand building or performance marketing objectives if the selection is correlated with both upstream and downstream behavior. However, if the selection is correlated primarily with downstream behaviors, then finer filtering is wasteful for upstream performance; in that it may eliminate potentially viable audience targets. Thus choice of journey stage seeds as a function of marketing objectives is an empirical question.

Our second set of research questions address the advertiser's challenge of audience expansion to increase the reach of the acquisition campaign. Specifically, we investigate the interaction effect between the journey stage of seeds and Lookalike match rank (based on level of similarity with seeds). The choice of journey stage seed and match rank range involves an exploration-exploitation tradeoff. If the Lookalike targeting algorithm's match score (and rank) with the seed is highly predictive of the lookalike's desired behavior (e.g., clickthrough, donations), then an exploitation strategy focusing on high match rank is likely preferable. However, as the predictive accuracy of the match score with the seed declines, exploration over lower ranks may be fruitful. Hence, we ask the following questions: First, should the match rank choice vary by the journey stage of the seeds, i.e., upstream versus downstream journey stage seeding? Second, do the differential effects of seeding and match rank vary based on the desired advertising outcomes, i.e., for brand marketers seeking upstream outcomes such as clickthrough versus performance marketers seeking downstream outcomes such as donations? Finally, given that ad performance is lower among low match ranked lookalikes, we consider an approach to improve their responsiveness. Motivated by recent work (Summers, Smith, and Reczek 2016; Shin and Yu 2021) that suggests that consumers perceive greater relevance to the advertised product when they know that firms have targeted ads to them, we test whether making targeting explicit in ads improves performance. We conjecture that the gains in performance from making targeting explicit would be greater for lower match rank individuals, for whom the ad is less relevant relative to the higher match ranked individuals for whom the ads are intrinsically more relevant.

To address these research questions, we conducted a set of randomized field experiments using Lookalike targeting on Facebook for new donor acquisition. Specifically, we test the effectiveness of different Lookalike audience profiling strategies by comparing ad effectiveness across the corresponding Lookalike audience sets. We use Facebook's randomized audience experiment feature to ensure that each Lookalike audience set receives an equal chance of winning the bid for ad exposure.[2] With its size and scale of 2.6 billion active monthly users worldwide (Statista 2020), Facebook is an ideal platform for an initial study of questions around Lookalike targeting.

A key practical challenge in assessing the effectiveness of any targeted advertising strategy is

[2]Note that in contrast to A/B tests to assess creative elements of different ads, where similarity of audiences across ads is critical for the experiment to be valid (e.g., Braun and Schwartz 2021), our research questions on lookalike targeting depends on identifying different audiences for different seeding strategies.

to have an appropriate control against which to measure lift. Researchers in recent ad-lift studies have therefore partnered with advertising platforms to let their algorithm identify a target for an ad, randomize these potential targets into the treatment ad or a control (e.g., Johnson, Lewis, and Nubbemeyer 2017), and use the difference in outcomes to measure incremental lift. This issue is particularly important when studying advertising on potentially "warm" targets, who have already embarked on the journey and therefore likely to purchase even without ads. However, such platform-level cooperation is practically difficult to obtain for individual advertisers that rely on large ad platforms like Facebook. We therefore consider an alternative strategy. First, note that since Lookalike targeting is focused on "cold" customers, who have not yet embarked on the journey, the issue of customers buying without an ad is already less pronounced in the context of new customer acquisition. But to the extent consumer needs in the category arise exogenously or these are larger brands with a potential high baseline of interest, different seeding treatments may target lookalikes with differential propensity to purchase without ads. To address this concern, we conduct our advertising experiment with a relatively small firm/non-profit for whom there is typically no new customer sales or new donor donations in the absence of advertising. We demonstrate that the outcomes from a control group not exposed to the treatment ad is close to zero, and hence the outcomes of the treated group can be considered as lift.

Our key findings are as follows: Overall, despite the cold start nature of lookalike audiences, seeding on other's journey stages can improve advertising performance—third parties such as Facebook can indeed identify factors unobserved (to the advertiser) that move individuals along the journey and can be correlated with lookalikes. However, the gains from moving seeds along the journey differ by advertiser objective. For a performance marketer seeking downstream journey outcomes such as donors, donation rates increase as one moves further down the journey stage in seeding. In contrast, for a brand marketer seeking more upstream outcomes (e.g., clickthrough), going further down the journey stage does not improve outcomes.

Second, we find an interaction effect between upstream/downstream stage seeding and Lookalike match rank. For downstream stage seeding (has donated), ad effectiveness decreases significantly when match rank is reduced from top 1% to 1%-2%. However, when seeded in the upstream stage of the journey (visited website), there is little performance difference between the top 1% and the top 1-2% match rank ranges. Therefore we conclude that performance marketers seeking downstream outcomes should use an "exploitation" strategy by targeting their advertising on lookalikes that have the highest match with seeds (top 1%), while brand marketers can benefit from a more "exploratory" strategy by expanding their targeting to lower ranked matches (1%-2%).

Finally, we find that performance marketers can alleviate the sharp drop in targeting effectiveness with low match ranks by highlighting that the ads are targeted to them. But explicit targeting has little impact on the high match rank individuals. Overall, making targeting explicit makes donor acquisition cost comparable for high and low match ranks. As differences in match rank brings in different lookalike customers, the benefit from making targeting explicit is entirely additive for performance marketers across match rank ranges in terms of incremental customer acquisition.

The rest of the paper is organized as follows. After outlining the links to related literature and providing background on the Lookalike targeting problem, we describe the details on the field experiment setting. We then discuss the experiments and results associated with journey stage seeding and the moderating role of match rank respectively. We conclude with suggestions for future research.

RELATED LITERATURE

This paper is connected primarily to the literature on targeting in advertising, and especially digital advertising. Table 1 provides an overview of how Lookalike targeting differs from targeting methods that have been studied in the literature. Digital channels have facilitated many novel ways of targeting; as digital traces left by consumers reveal contextual and real time information about preferences, immediate intent, and their stage along the customer journey. We classify targeting strategies into two broad groups: (1) contextual advertising based on contemporaneous information, and (ii) those that are based on user history. Examples of contextual advertising include keyword search (e.g., Agarwal, Hosanagar, and Smith 2011, Rutz and Bucklin 2011), and mobile targeting that leverages location and time information (e.g., Luo et al. 2014, Danaher et al. 2015, Fong, Fang, and Luo 2015). Targeting based on user history include those based on very short run history along a customer purchase journey (e.g., retargeting as in Lambrecht and Tucker 2013) or those based on a longer history of consumer behavior, interests, and preferences (e.g., Trusov, Ma, and Jamal 2016).[3] In contrast to all of this literature which focuses on improving targeting effectiveness based on behaviors of the same targeted individual, Lookalike targeting uses information on a seed group of *other individuals* to target new prospects based on their similarity with the seeds.

Further, the above individual-level targeting—contextual and user profile based—literature is built around conversion of customers who have already embarked on the customer purchase journey. As they are closely tied to "when" the customer is likely to be interested in the product, the conversion rates from such targeting tend to be quite high. However, these techniques are not particularly amenable to new customer acquisition, where customer needs and interest has to be initially stirred by advertisers and are less likely to be interested "at the moment." The most common approach for customer acquisition tends to be traditional display or banner advertising—with targeting around desired demographics, interests and behaviors proxied by the content sites that they visit (Manchanda et al. 2006, Goldfarb and Tucker 2011, Neumann, Tucker, and Whitfield 2019). With third party providers having considerable access to user demographics, interests and behaviors, advertisers can now request desired user profiles for new customer acquisition, but there are concerns about their accuracy and effectiveness (e.g., Neumann, Tucker, and Whitfield 2019). Lookalike targeting is particularly useful for customer acquisition because it provides advertisers an opportunity to leverage on massive amounts of third party data to seek out new customers that are not currently in a firm's database, nor have recently taken actions that indicate interest in the brand or category. To the best of our knowledge, no prior research has investigated how advertisers

[3]Another relevant paper related to the effect of advertising along the purchase funnel in an offline supermarket setting is Seiler and Yao (2017). They show the differential impact of advertising based on the customer's position along the conversion funnel, but their focus is not on targeting.

Table 1: Relationship to Previous Literature on Targeting

Targeting strategies	Example studies	Target based on history of	Firm Objective	Target Criteria Pre-specified?
Keyword search	Agarwal, Hosanagar, and Smith (2011) Rutz and Bucklin (2011)	Focal	Convert during journey	Yes
Contextual based on journey	Luo et al. (2014) Danaher et al. (2015) Fong, Fang, and Luo (2015) Li et al. (2017) Ghose, Li, and Liu (2019)	Focal	Convert during journey	Yes
Past purchase/browsing	Rossi, McCulloch, and Allenby (1996) Pancras and Sudhir (2007) Rafieian and Yoganarasimhan (2020)	Focal	Convert based on history	Yes
Retargeting on recent browsing	Lambrecht and Tucker (2013) Sahni, Narayanan, and Kalyanam (2019) Jiang et al. (2021)	Focal	Convert during journey	Yes
Lookalike targeting	**Our paper**	**Others**	**Acquire customer without journey/purchase history**	**No**

can improve targeting efficiency by using information in the journey stage of *others*.

Moreover, targeting approaches typically require marketers to enumerate attributes and conditions of the target audience for acquisition. Lookalike targeting, on the other hand, does not require advertisers to pre-specify target behavioral/descriptive profiles; advertisers can simply provide a list of people with desired behaviors, and then let the third party match customers within their database and identify those with similar characteristics. This makes targeted advertising not only easier to implement in practice, but also can be used even by firms seeking to acquire new customers but don't have a large existing customer database. We also note that while there has been some prior research focused on engineering issues in designing Lookalike profiling and targeting algorithms to maximize ad performance (Liu et al. 2016; Popov and Iakovleva 2018; Cotta et al. 2019), there has been little work on studying the effectiveness of Lookalike targeting and managerial choices in "implementing" Lookalike targeting as has been done with various other digital advertising methods.

Finally, we note similarities and contrasts between seeding strategies in Lookalike targeting and the distinct literature on seeding strategies within networks. In the network literature, seeding strategies are focused on selecting the optimal set of individuals in the social network who, given their position in the network, are most likely to exert peer influence in spreading word of mouth (e.g., Aral and Walker 2012; Domingos and Richardson 2001; Hinz et al. 2011; Kumar and Sudhir 2019). Examples include targeting opinion leaders who are highly connected and located in the central hub (e.g., Goldenberg et al. 2009), or located in dense regions of the network (e.g., Kitsak et al. 2010). Others underscore the importance of accounting for homophily—correlated behaviors or similarities among neighboring individuals in the social network (McPherson, Smith-Lovin, and Cook 2001)—in measuring the effectiveness of seeding strategies (e.g., Aral, Muchnik, and Sundararajan 2013; Nejad, Amini, and Babakus 2015). On the other hand, seeding in Lookalike targeting is based on the idea that similar individuals profiled from a seed database are likely to behave in a similar manner desired by the firm. The key difference is that network based targeting relies on the observed choices of individuals to be close to each other on some chosen network dimension (geography, social connection etc). However, with Lookalike targeting there is no such observable choice in terms of relationships—we are merely relying on a correlation, which may be causally rooted in an underlying set of latent variables–and through unobservable selection in who moves through stages of the journey. Thus the effectiveness of Lookalike targeting has to be empirically assessed based on whether there are effective latent variables proxied in third party data that causally drive similarities between seeds and lookalikes.

BACKGROUND

In this section, we provide relevant background on: (1) the Lookalike targeting problem, (2) the Lookalike targeting platform on Facebook, and (3) the empirical context in which we conduct our field experiments.

The Lookalike Targeting Problem

Advertisers have to make two key choices that determine Lookalike targeting success: (i) a seed set to perform Lookalike (similarity) modeling, and (ii) a Lookalike match set (i.e., match rank range) from which the Lookalike targeting algorithm will determine ad targets through adaptive learning.[4] Given the large pool of potential Lookalike candidates in the ad platform (potentially in the tens of millions of users), performance can greatly depend on declaring the right Lookalike audience set for ad targeting.

We formalize the problem as follows. The search space for Lookalike targeting is determined by two variables: the initial seed set $s \in S$ and match set $m \in M$, where M is the universal set of possible choices for the match rank range provided by the third party provider to the advertiser. The space of seed sets S that an advertiser can choose includes all possible seed sets available to the advertiser, both from its own CRM database and the seed options provided by the third party platform. Targeting performance depends on the strength of unobserved (to advertiser) correlation in desired behavior between seeds and matching lookalikes in the match set. The possible seeding

[4] Within the pre-defined lookalike audience pool, dynamically optimizing the exploration-exploitation tradeoff can save advertising costs (e.g., Schwartz, Bradlow, and Fader 2017) and improve donor acquisition.

strategy is infinitely large, given the large degrees of freedom coming from choosing the source and size of the seed sample. The third party ad platform typically provides a (finite) set of options for the choice set M from which m can be specified by the advertiser. Facebook, for example, allows up to 10% lookalike match rank range (in one percentile increments) in terms of its internal similarity score between the seeds and the lookalikes.

It is important to understand that the pair (s, m) jointly determines the quality of the Lookalike audience set, $L_{(s,m)}$, the candidates for ad targeting. In other words, advertisers determine $L_{(s,m)}$, the search space for the third party Lookalike targeting algorithm, denoted by a, to perform adaptive learning and show the ad to the targeted set of $l \in L_{(s,m)}$ individuals who the algorithm selects to explore and exploit.[5] Hence, the interaction between the unobserved correlations in behavior and potential noise present in the seeds s and match set m influence the exploration-exploitation tradeoff in effective ad targeting.

The advertiser's challenge is to identify the best seed and match set policy π, from among the infinitely large set of possibilities for $L_{(s,m)}$. While the size of seed set $|s|$ is usually in the range of thousands, the size of Lookalike audience $|L_{(s,m)}|$ is usually in the millions. We define $l_{optimal}$ as the set of best (i.e., in terms of the likelihood of maximizing the campaign objective) target individuals that reside in the third party platform. The dispersion and characteristics of $l_{optimal}$ individuals in the L sample space are unknown to the advertiser. The advertiser's choices boils down to specifying the search space by selecting the match rank range (m) that maximizes the likelihood of containing $l_{optimal} \in L$, given a (which is exogenous to the advertiser). Hence, the optimal seed-match policy that maximizes the campaign objective can be written as follows:

$$\pi := \underset{s \in S, m \in M}{\arg\max} P(l_{optimal} \in L_{(s,m)} | y, a).$$

Our primary research questions focus on generating insights around how an advertiser should choose s and m. Further, we recognize that the effectiveness of lookalike targeting may be moder-

[5]For example, Facebook ad campaigns first enter the learning phase in which the algorithm learns about the best audience set for ad delivery. For details, see https://www.facebook.com/business/help/112167992830700?id=561906377587030

ated by the content of the ads. While this moderation can kickstart a rich research agenda, in this paper, we explore a particular issue as it relates to our primary research questions: how making ad targeting explicit in the advertising content by highlighting similarity with others moderates the effect of s and m.

Lookalike Targeting Platform

As discussed earlier, many firms like Facebook, Google, Twitter and LinkedIn offer Lookalike targeting services. While the general approach to Lookalike targeting is similar across these platforms, we provide some background on Facebook Lookalike Audiences and the specific Lookalike targeting problem on which we conduct our experiments.

Advertisers that seek to use Lookalike targeting can either use proprietary seed data or seeds available through Facebook (e.g., Facebook page engagement). Facebook then uses these individuals as seeds to create an audience of individuals who are similar to, or "lookalike" to the seeds. If the advertiser wants to use proprietary seed data, it can upload the information about the seeds through the Facebook platform. Typically, an advertiser provides individual information about the seeds such as name, date of birth, gender, city, country, and email/phone numbers as personal identifiers. Facebook will then use the proprietary donor seed data to identify the matching profiles in Facebook through a secure hashing process. After the Lookalike audience generation process, Facebook deletes all uploaded information.[6]

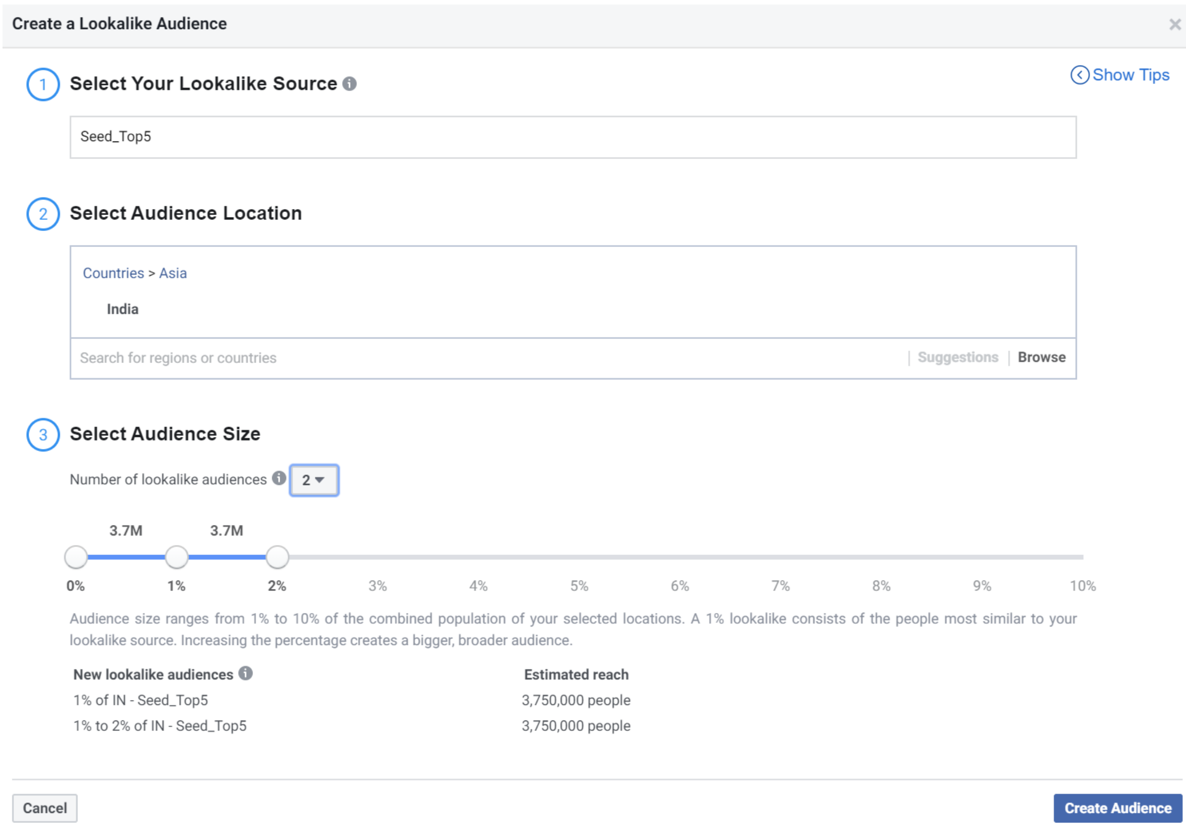
Facebook generally recommends a seed size between 1,000 to 50,000 to ensure good Lookalike matching. Then, advertisers decide the match rank and size a Lookalike audience from an interface shown in Figure 1. In our context, for example, requesting for top 0-1% Lookalike audience in India comprises of around 3.7M individuals on Facebook platform.[7] Similarly, requesting for the next match rank range of top 1-2% gives the subsequent 3.7M individuals ranked in Lookalike similarity. Creating a coarser audience diversifies the potential reach for new customer acquisition but reduces the algorithmically computed level of similarity between the Lookalike audience and

[6] https://www.facebook.com/business/help/112061095610075?id=2469097953376494

[7] Actual audience size would vary from 3.4-3.7M, depending on the active user condition at the time.

the seed set.

Figure 1: Lookalike Audience Generation



The Facebook platform allows advertisers to conduct ‘A/B/n’ audience split test experiments based on different seed set and various types of match rank range with respect to the seeds to search for Lookalike individuals on Facebook. Facebook allows up to five treatments in any test, i.e., $n \leq 5$. Facebook’s A/B test feature enables advertisers to compare performance between different audience selection strategies by ensuring that each treatment condition has equal chance of winning the bid using the lowest cost bid strategy, given the same ad budget.[8] For further details on the experiment setup and timeline of the ad campaigns, refer to Appendix B.

The Advertiser and Empirical Context

We conduct the research in partnership with an advertising partner, HelpAge India, a leading Indian non-profit organization that provides charitable support for the elderly in India. The non-profit

[8] https://www.facebook.com/business/help/1738164643098669

adopted Facebook Pixel technology to identify seeds on who had visited its website and donation page. In addition, Facebook provides seeding on individuals who had engaged with HelpAge India's Facebook page in terms of either likes or sharing of content on the page.

In addition to seeding using third party data, which is based on stages of the journey (up to website donation), the nonprofit advertiser also has a first party donor history database that can be used for seeding. For this lookalike targeting campaign, the organization used a monthly history of individual donor giving from April 2016 to February 2020 to segment consumers. The data included individual demographics such as name, date of birth, gender, city, country, and email that were uploaded on Facebook to be used as personal identifiers for seeding.

Table 2: Seed Data Statistics

	RFM Score	Recency (Month)	Total Donation Frequency	Total Donation Amount ($)	N
Top 5%	554.95	6.23	5.23	611.99	1829
	(0.22)	(4.41)	(5.98)	(2166.73)	
Top 10%	543.32	6.78	3.55	387.10	3659
	(17.65)	(4.58)	(4.67)	(1689.56)	
Total	356.96	24.19	1.33	89.34	36595
	(140.16)	(13.21)	(1.77)	(565.59)	

Note: Standard deviation in parenthesis

The CRM system at the nonprofit classified the firm's existing donors using RFM—based on prior donation behavior in the order of recency, frequency, and monetary value quintiles. To assess how donor quality differentiation impacted seeding, we used the RFM metric, as the organization felt that insights based on such donor differentiation can be valuable to them as these metrics will be available on an ongoing basis for their targeting campaigns. In unreported analysis, we indeed found that recency had the highest predictive power for predicting probability of donation, followed by frequency, and monetary value.[9] Table 2 details how the RFM descriptive statistics differ between top 5% and top 10% of donors. Note that selecting a larger seed base increases the

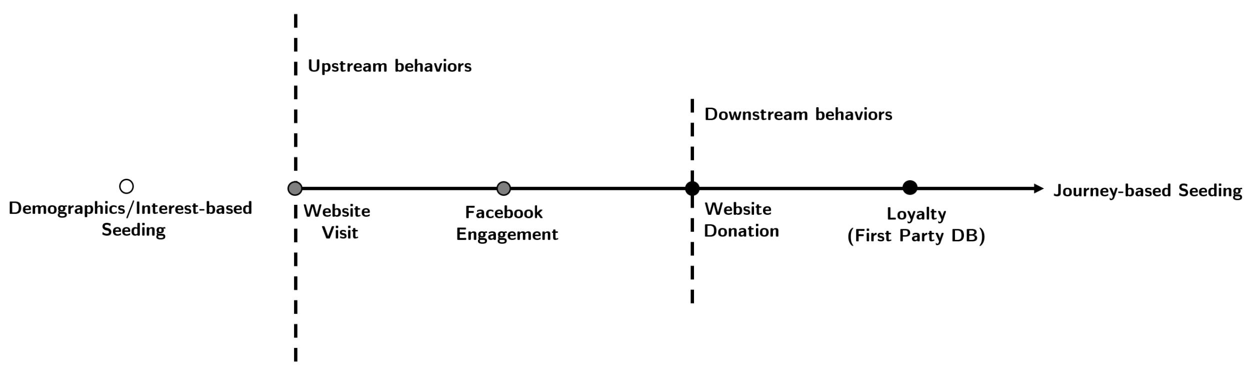
[9] Although many alternative techniques have been proposed to estimate customer value (e.g., Fader, Hardie, and Lee 2005; Zhang, Bradlow, and Small 2014), RFM remains widely used in industry for its ease of use and minimal data requirements.

potential sample size for Facebook profile identification, but reduces seed data quality. Doubling the seed size from the top 5% to top 10% led to a 79% increase in the standard deviation of RFM scores.

SEEDING BY JOURNEY STAGE

In this section, we address questions related to the advertiser's problem of determining the seed set based on journey stage information. We also consider how this seed choice should be moderated by desired ad outcomes; specifically, an upstream outcome such as clickthrough to the advertiser's website or a downstream outcome such as purchase/donation.

Figure 2: Seeding by Customer Journey Stage



To fix ideas, Figure 2 displays a schematic of a donor journey with respect to the focal firm as a funnel. The funnel shape indicates that only a fraction of individuals move from one journey stage to the next and there is potential selection along the journey. We label specifically the journey stages for which seeding data is available. The advertiser can seed based on various journey stage behaviors: (i) visited the charity website, (ii) engaged with charity social media (Facebook) site, and (iii) made a donation through charity website. Further, the advertiser has data on repeat donations through the charity's CRM system, which can be used to segment further and seed on levels of loyalty. Finally, as a benchmark for comparison, we consider targeting based on demographic/interest-based lookalikes without using journey data. We place this data on the left of the journey in the figure to indicate such demographic/interest-based targeting is available before seeds have embarked on journeys relevant to the target firm.

We label the initial visit and social media engagement behaviors as upstream behaviors, and the donation and loyalty behaviors (based on donation history) as downstream behaviors. This maps to the conventional nomenclature where desired behaviors for brand building (awareness, liking) are considered upstream behaviors, while sales/fundraising performance related behaviors (e.g., purchase or donation) are considered downstream behaviors. Based on data on existing customers at each journey stage, we create target Lookalike audiences of new individuals for whom, regardless of the journey-based seeding strategy, have not yet embarked on a journey with the focal firm, but have a higher likelihood of desired response.

Research Questions

Our first research question is to assess if seeding based on the journey stage of *another individual* is effective in targeted new customer acquisition. As discussed, it is an empirical question as to whether greater matching in terms of variables available in the third party database between the seeds and targets will lead to similar advertiser desired behaviors seen among the seeds for the lookalikes; that is, if a seed has clicked through or donated, it is not a priori obvious that a highly matched lookalike is also likely to engage in the same behavior.

The answer depends on whether there are some variables captured by the third-party on seeds and lookalikes that causally lead to the desired advertiser behavior. For this to happen, there should be selection in who moves along the customer journey, and the third party's database should possess variables that proxy for the selection and use them in the Lookalike targeting algorithm. If, however, reaching further down the journey stage is based mostly on brand-specific, contextual, or transient factors, it is less likely that similarities between seeds and lookalikes alone would be a useful targeting predictor, because lookalikes also need to be in similar contextual and transient factors. In other words, there needs to be brand-invariant unobserved (to advertiser) factors that correlate the seed's characteristics inherent to the journey stage with the lookalike's propensity of advertiser desired journey behaviors. Finally, given that typically third party databases and

Lookalike matching algorithms are blackboxes to advertisers (and researchers), we also assess whether journey-based seeding itself is effective at all, compared to other relevant benchmarks.

If indeed, there is evidence that Lookalike targeting effectiveness increases as one moves downstream along the journey, we then assess a second set of research questions. These questions are related to the managerial problem of whether the choice of seeding along the journey stage is moderated by the advertiser's objectives—brand versus performance marketing. The former needs to optimize on upstream behaviors (proxied by clicks), while the latter needs to optimize on downstream behaviors (proxied by purchase/donation). For an advertiser seeking an upstream outcome such as click-through for an ad to facilitate brand building, would it be sufficient to seed on individuals who had previously shown interest in the firm by visiting its website? Previous advertising studies show mixed results on the role of user activity mode and purchase journey stage in improving upstream advertising performance. Danaher and Mullarkey (2003) showed that display banner ads have more influence on users on exploratory browsing mode than those who are goal-directed toward a particular task. Hoban and Bucklin (2015) assessed ad effectiveness across different stages of the purchase funnel and showed that display ads had a positive effect on website visits for three of the four purchase funnel stages (i.e., nonvisitor, converted user, and authenticated user), but not on previous visitors who did not create an account. These results inform the need for optimizing journey-based seeding strategies depending on the firm's campaign objective for acquisition, whether it is to raise initial brand awareness or drive conversions.

Further, our study develops ideas closely related to previous work on ad effectiveness and advertising objectives, yet different in more than nuance in approach to using the selection effects in the purchase journey as seeds. New issues arise when creating lookalikes based on journey of existing consumers for acquisition. For example, selection imposes a cost to the advertiser in that it narrows the size of the potential seed set with whom the advertiser can generate the lookalikes. By the same token, if an advertiser seeks downstream outcomes (e.g., purchase, donation), would it be sufficient to only seed on individuals who have donated once in the past? Or would there be incremental value in seeding based on further downstream behaviors along the journey (e.g., loyalty level, lifetime value, WOM)?

Experiment Design

Our study focuses on a cold prospecting problem of picking the right audience to target. This means the evaluation of ad performance across different audiences is critical. We thus leverage Facebook's randomized audience experiment feature to assess the performance of different seeding strategies. It ensures that each audience treatment set has an equal chance of winning the bid.

Given our primary research focus around journey based seeding effectiveness, we consider four types of seeds that reflect different stages down the customer journey: upstream journey seeding on (i) Website Visits and (ii) Facebook Page Engagements and downstream journey seeding on (iii) Website Donations and (iv) Customer Value using First Party Purchase History Data. For seeding based on customer value, we use the top 5% of customers based on RFM from the donor database of HelpAge. Finally, since HelpAge has been using demographic targeting for new donor acquisition and wanted to assess the incremental value of Lookalike targeting relative to its current practice, we used demographic targeting as the fifth treatment baseline. We used the same criterion used by HelpAge to acquire a prospect list for offline targeting. This treatment does not involve seeds, but advertisers can specify desired demographic profile to target.[10] Table 3 presents the audience split testing experiment design with the five treatments. For all treatments, we keep the match rank range identical— top 1%, and allocated the same ad budget. We note that the size of the seed database is larger for more upstream journey stages due to the funneling nature of the selection that occurs as customers move over journey stages. Appendix B provides further details on the experiment setup.

Results

The outcomes for different Lookalike targeting treatments and the demographics benchmark are presented in Figure 3. First, we find that seeding based on journey stage of *others* improves Looka-

[10] Appendix A provides details on the firm's desired criteria for demographic targeting. As Facebook restricts audience testing feature to five treatments, our design exhausted the feasible number of treatments.

Table 3: Seeding by Journey Stage: Experiment Details

Journey	Seeding/Targeting Criterion	Lookalike Match	Seed Data Size	Target/Lookalike Audience Size
Upstream	Website Visit	Top 1%	75,000	3,300,000
	Facebook Page Engagement	Top 1%	32,000	3,900,000
Downstream	Website Donation	Top 1%	2,000	3,600,000
	Top 5% RFM	Top 1%	2,164	3,700,000
Baseline	Demographics	-	-	660,000

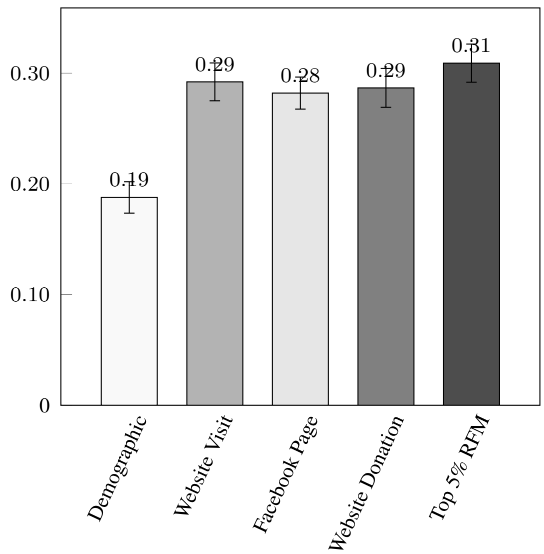
Note: 1) *Website Visit*: individuals who visited the firm's website in the past 120 days tracked by Facebook pixel.
2) *Facebook Page Engagement*: individuals who engaged in the firm's Facebook page in the past 120 days.
3) *Website Donation*: individuals who donated online in the past 120 days tracked by Facebook pixel.
4) *Demographics*: Details of demographic-based targeting presented in Appendix A.

like targeting performance, relative to the demographic-based benchmark. But to claim that these differences are due to effects of seeding by journey stage, it is important that the baseline visits (and therefore donations) without advertising are zero and not heterogeneous. As we note in the introduction, in our non-profit setting, new donor acquisition seldom occurs on their own without solicitation, and hence this assumption is likely to be valid. Nevertheless, we assessed this using an experiment described in Appendix C. Overall, the experiment confirms the validity of our assumption.

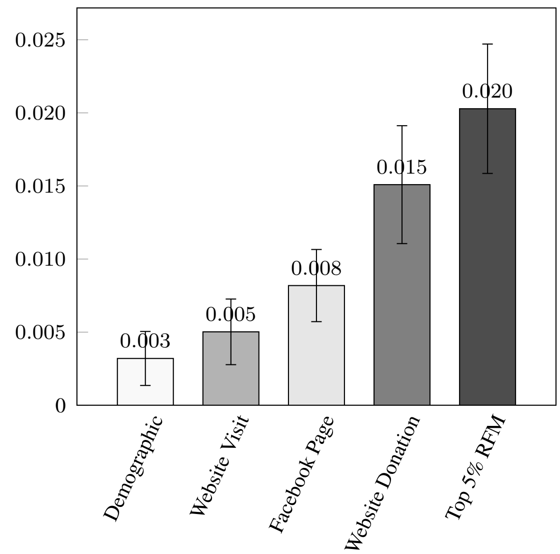
Furthermore, we find that the incremental benefit of additional movement down the journey only exists to the extent of the threshold behavioral correlation between the journey stage information and observed behavior. Figure 3a shows that indeed, the differences in clickthrough rates based on the various upstream (website visits and Facebook page engagement) and downstream journey seeds (website donations, top 5% RFM) are not statistically significant. Thus, our findings show that past the interest stage journey proxied by visits, the journey stage correlations that move customers down the journey do not add incremental value in improving the likelihood of upstream behaviors. Hence, advertisers with the goal of brand marketing need not use downstream journey data to seed Lookalike targeting.

We also find that in terms of downstream advertising goal of acquisition (donation), Lookalike targeting effectiveness increases down the journey stage seeding strategy up to seeding on web-

Figure 3: Seeding by Journey Stage



(a) Click-through Rate (%)



(b) Donation Rate (%)

Note: Error bars represent standard error.

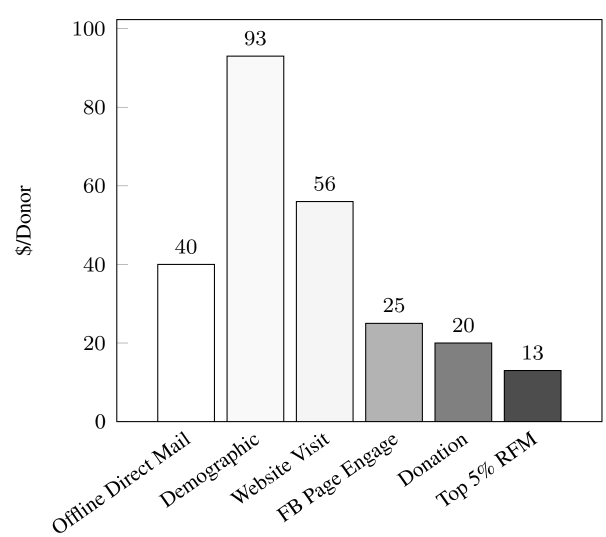
site donations. Figure 3b illustrates this monotonic improvement in donation rates for lookalikes. Specifically, the difference in donation rates from using website visits (upstream stage) to website donations (downstream stage) as seeds is statistically different ($p<0.05$). However, we did not find statistically significant differences between website donations and Top 5% RFM ($p=.49$). This finding is consistent with our conjecture that seed stage does not need to go further than the marketer desired outcome (donation).[11]

Overall, seeding on journey stage of *others* matters for Lookalike targeting, and journey stage has differential impact on upstream versus downstream behavioral outcomes of the targeted lookalikes. For upstream marketing, in which the marketer's objective is to increase interest (proxied by clicks), our results suggest that upstream journey stage seeds are sufficient, and investing on later journey seeds does not yield incremental value. For performance marketers, whose primary goal is to drive conversions/donations, there is a monotonic increase in donation rates as one moves further down the journey, but beyond the purchase journey stage, fine-tuning of seed quality using customer history and loyalty information adds limited incremental value.

[11] Additional unreported experiments with seeding based on different ranges of RFM confirmed this conclusion.

We note that seeding data using (i)-(iii) up to website donation is available through the third party, i.e., from Facebook through its Facebook Pixel (cookie-type) technology. Thus, any advertiser who has allowed the third party to include Facebook Pixel can seed for lookalike targeting without access to any of its own first party data. However, seeding using data on loyalty (repeat purchases and engagement across other channels etc.) can be done only with first party data. Given that not all advertisers have their own extensive customer lists and rich information on existing or prospective customers, comparing the relative benefits from HelpAge's first party data as seeds relative to seeds along the donor journey using third party data is of general interest to advertisers. This is especially relevant for young brands without extensive customer history data and whose initial goals are often to raise brand interest and upstream behaviors. Whether late stage and loyalty-based seeding adds value for targeting can provide practical guidance on data investment for seeding purposes. Further, given the increasing privacy challenges to the use of third party cookies, these questions about the value of investments in first party data also gain increased urgency and practical importance.

Figure 4: Cost Per Acquired Donor



Finally, along the line of increase in donation rate based on seeding along the customer journey, the cost per donor acquisition also falls (see Figure 4). In the graph, as a benchmark we include the average cost of traditional offline targeting campaigns by direct mail. Our results show that online Lookalike targeting is not always more cost-effective than offline targeting. Specifically, here we find that demographic targeting (as done using the variables used by this non-profit) and upstream seeding on website visit are less cost-effective than doing offline direct mail targeting, for performance marketers focused primarily on donor acquisition.

MODERATING EFFECT OF MATCH RANK ON JOURNEY STAGE SEEDING

Having demonstrated the effectiveness of using journey stage as seeds for Lookalike targeting, we next address the advertiser's choice of the second strategic variable, match rank range, and its interaction with journey stage seeding. The seed set and the match rank range specified by the advertiser jointly determines the "search space" for Facebook to look for targeting prospects.

Our second set of research questions addresses the advertiser's challenge of audience expansion to increase acquisition reach. As different match ranges brings in different prospects, when should advertisers explore lower match ranks, and when should they exploit higher match ranks? To the extent the predictive accuracy of the match scores is a function of the seed quality, we focus on the interaction between journey stage seeding and match rank range. First, how sensitive is ad performance to the reduction in match rank for upstream journey stage seeding versus downstream journey stage seeding? Second, how do these effects vary based on desired advertising outcomes (i.e, brand vs. performance marketing)?

We expect less information content in upstream journey stage seeds, and therefore lower predictive power of seed-lookalike match score in predicting desired targeted behaviors. We therefore conjecture that it should be optimal to explore the space of lookalikes more by experimenting with a wider range of ranks when seeding by upstream journey stages. But as one moves along the journey, selection should lead to more information in downstream journey stage seeds and therefore higher accuracy for match score between seeds and lookalikes in predicting desired behaviors. So we conjecture that it would be optimal for advertisers using downstream journey stage seeding to be more conservative in exploring for lookalikes and exploit the information embedded through selection in these seeds by focusing on a narrow range of highest ranked matches.

Further, we conjecture that the gains from exploration by searching among lower ranked lookalikes will be smaller for downstream behaviors (e.g., purchases, donations) than for upstream behaviors (e.g., clicks). This is because marketing campaigns for downstream behaviors need more focused targeting, and given the inherent selection and information embedded in the movement down the journey, while upstream behaviors may be induced by more transient and contextual factors. We therefore conjecture that it would be optimal for advertisers to choose a more narrow range of highest ranked matches when targeting for downstream behaviors, but explore more among lower ranked matches when targeting for upstream behaviors.

Experiment Design

We conduct two sets of randomized audience experiments to assess how the effects of high match rank (top 0-1%, exploitation of seed-lookalike correlation) and lower rank (top 1-2%, exploration of different lookalike search space) are moderated by upstream and downstream journey seeding strategy, as shown in Table 4 and 5.

For upstream journey stage seeding, we use (i) Website Visits and (ii) Facebook Page Engagement. For downstream stage journey stage seeding, we use the high value donors (in terms of RFM) present in the nonprofit's CRM data. Along with this, we further assess the robustness of downstream stage seeding effectiveness by diluting the first party seed quality from highest value (top 5% RFM) to top 10%, doubling the seed size.[12] Finally, similar to our first experiment on seeding, same ad budget is allocated for each ad set, and Facebook's audience A/B test feature ensures each Lookalike audience set has an equal chance of winning the bid.

[12]To be consistent with our first experiment on journey stage seeding, we wanted to conduct the experiment with website donations and Top 5% RFM to assess downstream journey stage seeding. But the non-profit executives were keen on understanding differences arising from loyalty information the CRM data (Top 5% vs Top 10% RFM).

Table 4: Match Rank Range: Interaction with Upstream Journey Stage Seeds

Journey Stage	Seed Data	Match Rank Range	Seed Data Size	Lookalike Audience Size
Upstream	Website Visit	Top 0-1%	46,000	4,100,000
	Website Visit	Top 1-2%		3,900,000
	Facebook Engagement	Top 0-1%	15,000	4,000,000
	Facebook Engagement	Top 1-2%		4,000,000

Note: 1) *Website Visit*: individuals who visited the firm's website in the past 120 days tracked by Facebook pixel.
2) *Facebook Page Engagement*: individuals who engaged in the firm's Facebook page in the past 120 days tracked by Facebook pixel.

Table 5: Match Rank Range: Interaction with Downstream Journey Stage Seeds

Journey Stage	Seed Data	Match Rank Range	Seed Data Size	Lookalike Audience Size
Downstream	Top 5% RFM	Top 0-1%	1829	3.7-3.9M
	Top 5% RFM	Top 1-2%		
	Top 10% RFM	Top 0-1%	3659	
	Top 10% RFM	Top 1-2%		

Note: During this experiment period, Facebook did not provide the exact audience size, but provided a range with the following message: "To protect the privacy of people on our platforms, we aren't showing the audience size." The suggested range for Lookalike Audience size is between 3.7-3.9 million, 1% of the population of Facebook users in India at the time of the experiment.

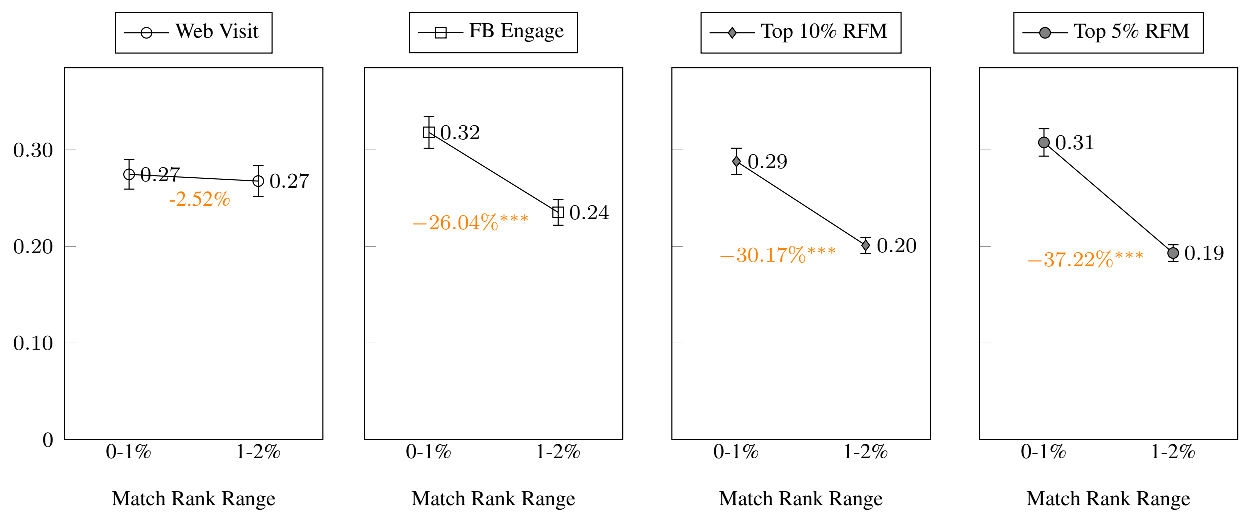
Results

The interaction effects of match rank range and journey stage seeding are presented in Figure 5a and 5b. First, we show that the role of higher match rank range become more critical going down the journey stage seeding strategies (i.e., from website visits and Facebook page engagement to latest stage of "having donated" history). That is, the performance gap between the highest 1% and subsequent 1-2% match rank widens going down the journey stages of the seeds, both in terms of clicks and donations. When seeded on upstream stages of the journey, we find that ad performance is not very sensitive to the choice of match rank range and the clickthrough rates and donation rates are not statistically significantly different for web visits and FB engage seeding. This suggests that the "optimal" (i.e., in terms of ad campaign objective) lookalike candidates may be more dispersed over the different match rank range in upstream stage seeding, when the behavioral correlation between seed and lookalike is more noisy. Hence, the match rank range decision becomes less important under upstream stage seeding, and advertisers can explore "suboptimal" (in terms of third party match rank range criterion) lookalike search spaces. That is, the cost of exploration for

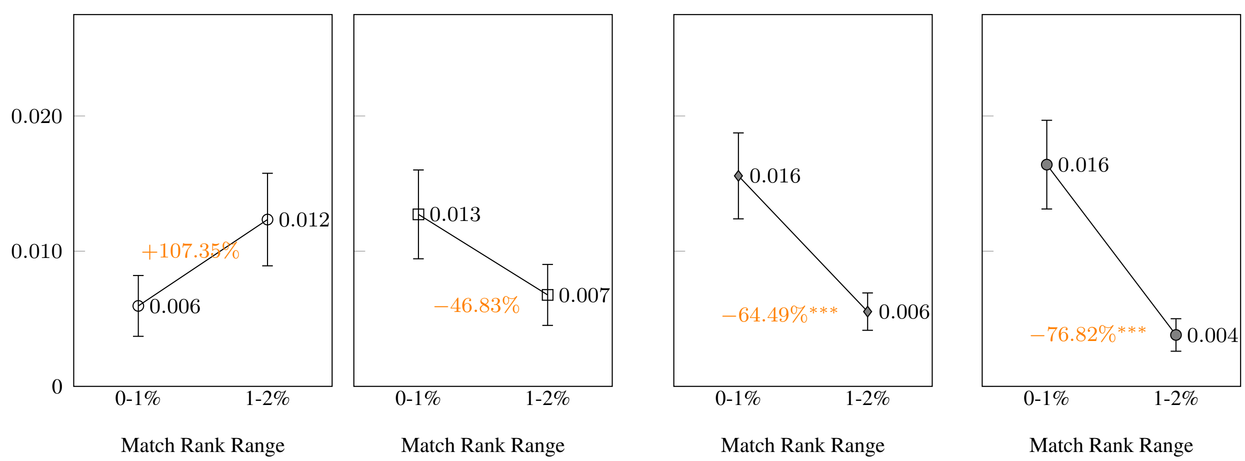
Lookalike audience expansion is lower for upstream seeds.

Figure 5: Moderating Role of Match Rank on Journey Stage Seeding

(a) Clickthrough Rate (%)



(b) Donation Rate (%)



Note: * p<0.05, ** p<0.01, *** p<0.001. Error bars represent standard error.

When seeded on downstream stages of the journey, however, advertising performance decreases significantly when match accuracy is reduced from top 0-1% to the next 1-2% of match rank. Reduction of match rank resulted in 37% (30%) reduction in click-through rates and 77% (64%) drop in donations of Lookalikes of top 5% (10%) seed individuals. Second, ad performance differences due to reducing seed quality by doubling the seed sample size from top 5% to 10% of donors seems negligible. While there appears a slight reduction in outcomes for both clicks and donations, the differences are not statistically significant. Hence, we find that clicks and donations are not very sensitive to seed audience quality reduction but can sharply decline with lower match rank. This suggests that when seeding on downstream stages of the journey, exploitation of match value by focusing on the highest rank works better, and ad performance relies critically on the third party audience profiling quality (in terms of ability to predict desired behaviors by lookalikes), relative to the quality of the first party seed data.[13]

We also find as expected, that match rank range is more critical for downstream behaviors (i.e., donation) than upstream behaviors (i.e., clicks). Moving from the top 1% to the top 1-2% match rank range resulted in a larger percentage decrease in donations compared to clicks for any journey based seeding strategy. Thus it is more challenging for performance marketers to expand customer acquisition by relaxing the match rank criteria than for brand marketers.

INCREASING TARGETING RELEVANCE FOR LOW MATCH RANK

In the last section, we find a large reduction in ad performance for low match rank especially when seeding with downstream journey stage. A natural explanation could be that low match rank individuals perceive the ad as less relevant to them. We therefore explore how performance marketers can increase acquisition reach while retaining performance when using downstream seeding strategy.

Recent literature in consumer behavior (Summers, Smith, and Reczek 2016) and analytical modeling (Shin and Yu 2021) suggest that consumers perceive greater relevance when they know that they have been chosen to be shown the ad based on the advertiser's data and algorithm. Digital ad platforms often embed messages in advertisements that explicitly let individuals know they are viewing a targeted ad or a recommendation. Examples include Amazon's "Recommended for you" product section and Netflix's "Because you watched..." content recommendation section. We therefore investigate whether making targeting explicit in an ad can improve ad effectiveness. We hypothesize that this improvement would be stronger for the lower match ranked lookalikes, who

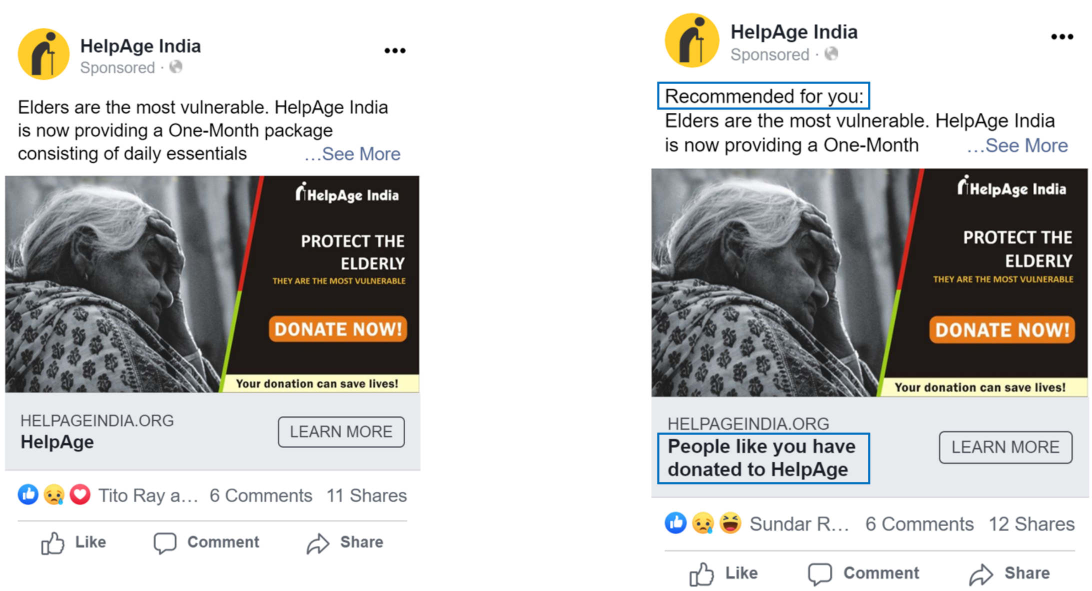
[13] We conducted an additional experiment on seed data quality, where we considered seed quality to be Top 6-10%. The clickthrough and donation rates continued to remain insensitive to the change in seed quality.

have lower natural fit of the message relative to higher match rank lookalikes.

Experiment Design

To increase the perceived relevance of targeted advertising, we make Lookalike targeting explicit by stating in the ad that the targets (lookalikes seeded on donation) were being shown the ad because *others like them* have donated to the nonprofit. Using a randomized field experiment design, Lookalike audiences of different match rank range (top 1% and 1-2%) were randomly allocated between treatment and control ad condition, as shown in Figure 6. For the controls, we did not mention anything about targeting.

Figure 6: Sample Screenshots of Experiment Design



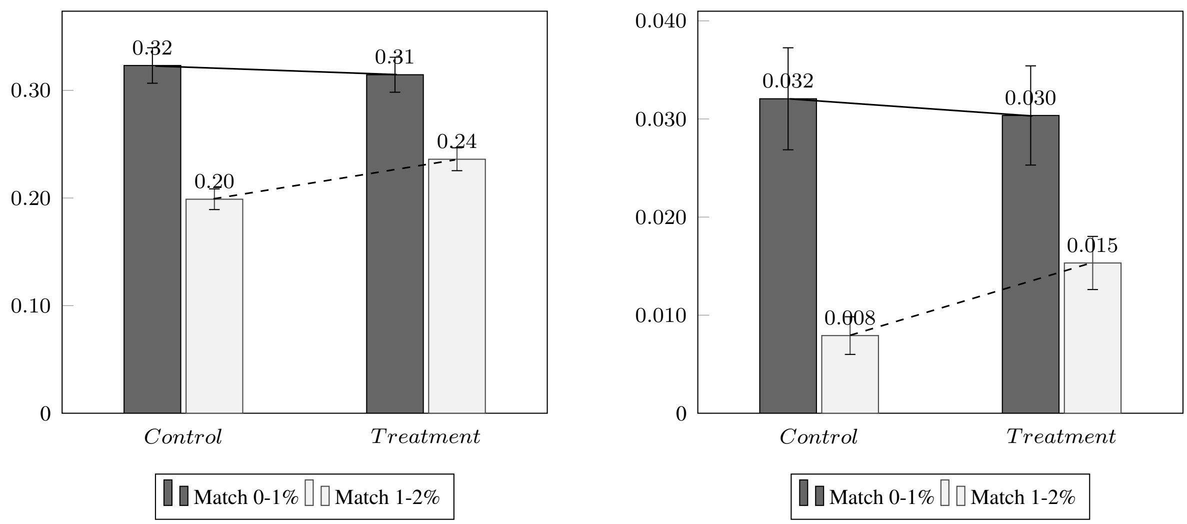
(a) Control: Targeting Implicit

(b) Treatment: Targeting Explicit

Results

Figure 7 illustrates how highlighting targeting differentially impacted those with high and low match rank. As hypothesized, for low match rank range (1-2%), we find that making targeting explicit improves both clickthrough and donation rates. Specifically, making targeting explicit improved clicks by 19% ($p < 0.01$), and donations by 93% ($p < 0.05$). However there is no significant difference for the high match rank range (0-1%). Overall, our results show that marketers can enhance ad performance by making Lookalike targeting appear more relevant when they use lower match rank ranges.

Figure 7: Interaction Effects of Explicit Targeting Message on Match Rank Range



(a) Click-through Rate (%) **(b)** Donation Rate (%)

Note: Error bars represent standard error.

Cost-Benefit Analysis

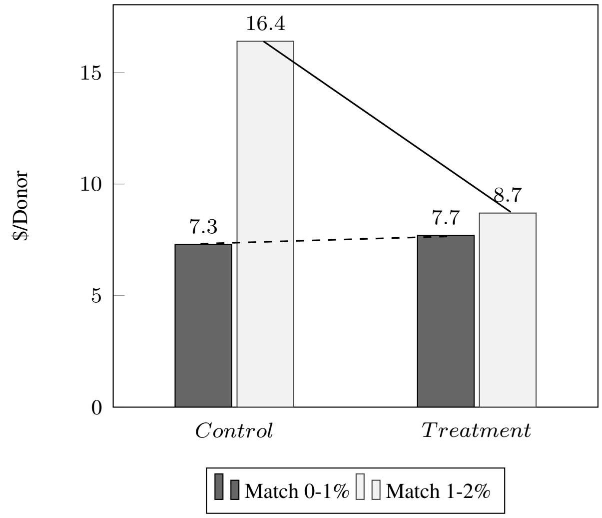
Table 6 presents the cost-benefit analysis of making targeting explicit. First, as expected from previous results, the cost per thousand impressions are indeed much lower for lower match rank in both the control and treatment conditions. Cost per click also drops with lower match rank.

Table 6: Advertising Costs: Match Rank and Explicit Targeting Message

Treatment	Match Rank Range	CPM ($) (Cost per 1,000 Impressions)	CPC ($) (Cost per Link Click)	CAC ($) (Cost per Acquisition)
Control	0-1%	2.35	0.73	7.33
Control	1-2%	1.30	0.65	16.40
Treatment	0-1%	2.35	0.75	7.74
Treatment	1-2%	1.33	0.56	8.71

Note: Downstream journey stage seeding (previous donors) was used to generate lookalikes.

Figure 8: Cost of Acquisition: Match Rank and Explicit Targeting Message



What is particularly interesting is how explicit targeting message interacts with match ranks on the cost of acquisition (see Figure 8). Even though the cost per impression is lower with low match rank range (1-2%), the cost per acquisition of customers is more than double ($16.40 versus $7.33) when the advertiser simply targets, but does not make the targeting explicit. In contrast, once targeting is made explicit, the cost of acquisition is much more comparable ($8.71 versus $7.74). This means that one can expand customer acquisition at comparable acquisition costs by highlighting the targeting when the lookalike match rank is lower (top 1-2%), but not making targeting explicit for higher match rank. And since differences in match rank brings in different lookalike customers, pursuing such a segmented messaging strategy for different match rank ranges is entirely additive (across two separate segments) for new customer acquisition. This is a managerially useful insight.

CONCLUSION

Lookalike targeting has emerged as an important ad targeting offering on most major digital advertising platforms. Unlike much of the focus in digital targeting based on one's own behaviors, Lookalike targeting is based on similarity in behaviors (and descriptors) with "seeds" chosen by the advertisers. While engineering aspects of designing effective Lookalike profiling algorithms had been explored in the fields of computer science and engineering, there is little academic research in marketing from the perspective of the advertiser. To the best of our knowledge, this is the first paper to empirically test the effectiveness of Lookalike targeting and provide guidelines for how brand marketers who seek upstream journey outcomes (e.g., clickthrough) and performance marketers who seek downstream journey outcomes (e.g., sales, donations) can effectively use Lookalike targeting. The paper focused on two critical choices faced by advertisers in Lookalike targeting: (i) seeding based on journey stage and (ii) seed-lookalike match rank. We also assess how making Lookalike targeting explicit impacts advertising effectiveness as a function of the match rank.

Overall, Lookalike targeting using other's journey stages can be effective; advertisers can exploit the information embedded through selection in the journey stages of others, captured in third party data to improve customer acquisition. We highlight that though previous research has suggested that past purchase behaviors are the best predictors of future behaviors for the same person (e.g., Rossi, McCulloch, and Allenby 1996, Pancras and Sudhir 2007), the current paper demonstrates that similarities with others, i.e., seeds exhibiting desirable behaviors (movement along journey), can be good predictors for marketers seeking those desirable behaviors. Importantly, we also show that the value of the information embedded in the selection from using seeds further down the journey is only relevant for a performance marketer seeking similar downstream outcomes. For brand marketers, seeking upstream outcomes, there is little incremental value from the refined information through selection.

Upstream seeds typically have less information content than downstream seeds because upstream behaviors are more prevalent and there is no embedded information through selection. As such, match scores between lookalikes and seeds have less predictive power for desirable upstream behaviors than for downstream behaviors. Hence we find that it is best for performance marketers using downstream seeds and seeking downstream behaviors to "exploit" high match ranks, while it is more valuable for brand marketers using upstream seeds and seeking upstream behaviors to "explore" among lower ranked lookalikes. Performance marketers can cost-effectively expand their reach to lookalikes with low match ranks by making targeting explicit only to lower match-ranked lookalikes; with this segmented strategy, new donor acquisition cost becomes comparable for both high and low match rank lookalikes. Further, as high and low rank lookalikes belong to different segments, the benefit of this segmented messaging strategy is entirely additive.

Limitations and Future Research

We conclude with a discussion of limitations and suggestions for future work. First, research should consider generalizability and boundary conditions for Lookalike targeting across categories; specifically, are there categories where similarities among individuals who exhibit desired behaviors serve more or less effectively as a targeting tool? Also while we found final donation behavior to be most effective for targeting, are there potentially other sweet spots along the customer journey that could be practically effective for advertisers? For example, for products with a clear time component for purchase (e.g., tickets to sporting events), would consumers who begin the search journey for a particular event be more informative than past buyers of tickets to similar sporting events?

Second, while our targeting experiments were done using Facebook Lookalike Audiences, it would be useful and important to see how the results replicate across other platforms such as Google, Twitter and LinkedIn. Yet, given the scale and breadth of Facebook data, these results based on Facebook should be useful in its own right. If Lookalike targeting does not work with Facebook given its scale, it is unlikely to work for smaller platforms with less data and potentially less sophisticated targeting algorithms. We note that our focus on one platform is not dissimilar to studies of keyword search advertising that had focused on Google, even though there are other search engines (e.g., Bing). As such, the current initial evidence that Lookalike targeting works on Facebook is useful for scholars and practitioners, though future work should assess results on other platforms. Since the data available for Lookalike targeting can be different across platforms, framing research questions to test effectiveness based on the kind of data available to these platforms to inform when each of these platforms should be used can also be valuable. In that spirit, though the emphasis in this paper is based on Lookalike targeting with third party data, similar research questions could be considered with Lookalike targeting using second party data, where advertisers seek cooperation with specific firms who have data on consumption in related categories. It is possible that sensitivity to match rank may be lower with second party data in closely related categories.

Third, we measured the effectiveness of Lookalike targeting through the metric of immediate donor acquisition. Future research should explore longer-term effects in terms of ongoing donations/purchases and lifetime value. For example, even though we found that conditional on recent donations, donation rates are insensitive to seed loyalty (in terms of RFM), it is possible that seeding based on CLV may matter for acquiring longer-term higher CLV customers. While this issue may be less salient in a nonprofit donation setting, where all donations are incrementally valuable, in settings involving high acquisition and ongoing maintenance and retention costs, seed quality (in terms of CLV) may be more critical.

Fourth, we chose the experiment setting on donor acquisition at a nonprofit to minimize the measurement challenge over lift; it would be natural to study whether these results generalize to customer acquisition in for-profit settings. As a specific example, it is possible that individuals perceive the ad as less intrusive when coming from a nonprofit than from a for-profit.

Finally, there are some threats to the use of Lookalike targeting through third parties as browsers are increasingly making data collection through cookies challenging. To the extent that Facebook itself can identify purchases on the website through its Facebook Pixel, it is interesting in this setting that Lookalike targeting based on purchases does not require even uploads to Facebook of customer data for identifying seeds who have made donations. But as third party data collection becomes more challenging, it would be useful to consider alternative approaches to execute Lookalike targeting. For instance advertisers with first party seeding data, may need to partner/contract with specific second parties with the most relevant "lookalike" behavioral information to execute Lookalike targeting. Our work can serve as a useful empirical framework to assess the value of such partnerships and contracts. Overall, we hope our initial investigation into Lookalike targeting will be an impetus to explore a variety of related questions—given its extensive adoption across a range of digital platforms, and its critical importance for new customer acquisition.

REFERENCES

Agarwal, Ashish, Kartik Hosanagar, and Michael D. Smith (2011), "Location, location, location: An analysis of profitability of position in online advertising markets," *Journal of marketing research*, 48 (6), 1057–1073.

Aral, Sinan, Lev Muchnik, and Arun Sundararajan (2013), "Engineering social contagions: Optimal network seeding in the presence of homophily," *Network Science*, 1 (2), 125–153 Publisher: Cambridge University Press.

Aral, Sinan and Dylan Walker (2012), "Identifying influential and susceptible members of social networks," *Science*, 337 (6092), 337–341 Publisher: American Association for the Advancement of Science.

Braun, Michael and Eric M Schwartz (2021), "The A/B Test Deception: Divergent Delivery, Ad Response Heterogeneity, and Erroneous Inferences in Online Advertising Field Experiments," *SMU Cox School of Business Research Paper*, (21-10).

Cotta, Randell, Mingyang Hu, Dan Jiang, and Peizhou Liao "Off-Policy Evaluation of Probabilistic Identity Data in Lookalike Modeling," "Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining," pages 483–491 (2019).

Danaher, Peter J. and Guy W. Mullarkey (2003), "Factors affecting online advertising recall: A study of students," *Journal of advertising research*, 43 (3), 252–267 Publisher: Cambridge University Press.

Danaher, Peter J, Michael S Smith, Kulan Ranasinghe, and Tracey S Danaher (2015), "Where, when, and how long: Factors that influence the redemption of mobile phone coupons," *Journal of Marketing Research*, 52 (5), 710–725.

Domingos, Pedro and Matt Richardson "Mining the network value of customers," "Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining," pages 57–66 (2001).

Fader, Peter S., Bruce GS Hardie, and Ka Lok Lee (2005), "RFM and CLV: Using iso-value curves for customer base analysis," *Journal of marketing research*, 42 (4), 415–430.

Fong, Nathan M., Zheng Fang, and Xueming Luo (2015), "Geo-conquesting: Competitive locational targeting of mobile promotions," *Journal of Marketing Research*, 52 (5), 726–735 Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Ghose, Anindya, Beibei Li, and Siyuan Liu (2019), "Mobile targeting using customer trajectory patterns," *Management Science*, 65 (11), 5027–5049.

Goldenberg, Jacob, Sangman Han, Donald R. Lehmann, and Jae Weon Hong (2009), "The role of hubs in the adoption process," *Journal of marketing*, 73 (2), 1–13 Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Goldfarb, Avi and Catherine Tucker (2011), "Online display advertising: Targeting and obtrusiveness," *Marketing Science*, 30 (3), 389–404.

Hinz, Oliver, Bernd Skiera, Christian Barrot, and Jan U. Becker (2011), "Seeding strategies for viral marketing: An empirical comparison," *Journal of Marketing*, 75 (6), 55–71 Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Hoban, Paul R. and Randolph E. Bucklin (2015), "Effects of internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment," *Journal of Marketing Research*, 52 (3), 375–393 Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Jiang, Zhenling, Tat Chan, Hai Che, and Youwei Wang (2021), "Consumer Search and Purchase: An Empirical Investigation of Retargeting Based on Consumer Online Behaviors," *Marketing Science*.

Johnson, Garrett A, Randall A Lewis, and Elmar I Nubbemeyer (2017), "Ghost ads: Improving the economics of measuring online ad effectiveness," *Journal of Marketing Research*, 54 (6), 867–884.

Kitsak, Maksim, Lazaros K. Gallos, Shlomo Havlin, Fredrik Liljeros, Lev Muchnik, H. Eugene Stanley, and Hernán A. Makse (2010), "Identification of influential spreaders in complex networks," *Nature physics*, 6 (11), 888–893 Publisher: Nature Publishing Group.

Kumar, Vineet and K. Sudhir (2019), "Can friends seed more buzz and adoption?," Publisher: Cowles Foundation Discussion Paper.

Lambrecht, Anja and Catherine Tucker (2013), "When does retargeting work? Information specificity in online advertising," *Journal of Marketing research*, 50 (5), 561–576.

Li, Chenxi, Xueming Luo, Cheng Zhang, and Xiaoyi Wang (2017), "Sunny, rainy, and cloudy with a chance of mobile promotion effectiveness," *Marketing Science*, 36 (5), 762–779.

Liu, Haishan, David Pardoe, Kun Liu, Manoj Thakur, Frank Cao, and Chongzhe Li "Audience expansion for online social network advertising," "Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining," pages 165–174 (2016).

Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee Wei Phang (2014), "Mobile targeting," *Management Science*, 60 (7), 1738–1756 Publisher: INFORMS.

Manchanda, Puneet, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K. Chintagunta (2006), "The effect of banner advertising on internet purchasing," *Journal of Marketing Research*, 43 (1), 98–108 Publisher: SAGE Publications Sage CA: Los Angeles, CA.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual review of sociology*, 27 (1), 415–444 Publisher: Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA.

Nejad, Mohammad G., Mehdi Amini, and Emin Babakus (2015), "Success factors in product seeding: The role of homophily," *Journal of Retailing*, 91 (1), 68–88 Publisher: Elsevier.

Neumann, Nico, Catherine E. Tucker, and Timothy Whitfield (2019), "Frontiers: How effective is third-party consumer profiling? Evidence from field studies," *Marketing Science*, 38 (6), 918–926.

Pancras, Joseph and K. Sudhir (2007), "Optimal marketing strategies for a customer data intermediary," *Journal of Marketing research*, 44 (4), 560–578.

Popov, Artem and Daria Iakovleva (2018), "Adaptive look-alike targeting in social networks advertising," *Procedia computer science*, 136, 255–264.

Rafieian, Omid and Hema Yoganarasimhan (2020), "Targeting and privacy in mobile advertising," *Available at SSRN 3163806*.

Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby (1996), "The value of purchase history data in target marketing," *Marketing Science*, 15 (4), 321–340 Publisher: INFORMS.

Rutz, Oliver J. and Randolph E. Bucklin (2011), "From generic to branded: A model of spillover in paid search advertising," *Journal of Marketing Research*, 48 (1), 87–102 Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Sahni, Navdeep S, Sridhar Narayanan, and Kirthi Kalyanam (2019), "An experimental investigation of the effects of retargeted advertising: The role of frequency and timing," *Journal of Marketing Research*, 56 (3), 401–418.

Schwartz, Eric M, Eric T Bradlow, and Peter S Fader (2017), "Customer acquisition via display advertising using multi-armed bandit experiments," *Marketing Science*, 36 (4), 500–522.

Seiler, Stephan and Song Yao (2017), "The impact of advertising along the conversion funnel," *Quantitative Marketing and Economics*, 15 (3), 241–278 Publisher: Springer.

Shin, Jiwoong and Jungju Yu (2021), "Targeted advertising and consumer inference," *Marketing Science*, 40 (5), 900–922.

Statista "Number of monthly active Facebook users worldwide as of 1st quarter 2020," (2020) https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/, (Accessed July 17, 2020).

Summers, Christopher A., Robert W. Smith, and Rebecca Walker Reczek (2016), "An audience of one: Behaviorally targeted ads as implied social labels," *Journal of Consumer Research*, 43 (1), 156–178.

Trusov, Michael, Liye Ma, and Zainab Jamal (2016), "Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting," *Marketing Science*, 35 (3), 405–426.

Zhang, Yao, Eric T. Bradlow, and Dylan S. Small (2014), "Predicting customer value using clumpiness: From RFM to RFMC," *Marketing Science*, 34 (2), 195–208.

APPENDIX

A. Firm-provided Demographic Targeting Criterion

Prior to this study, the nonprofit has been using the following demographic criterion to request and purchase mailing lists from the third-party data brokers for new donor acquisition. We hence use the desired demographic specification as benchmark to assess the incremental value of Lookalike targeting. We outline the demographic criterion as below:

The Nonprofit's Cold Mailing List Criterion:

1. High networth working individuals: with at least ₹500,000 in investments in the stock market or mutual funds. Also, the annual income needs to be at least ₹500,000 and preferably ₹1,000,000 or more.
2. Region: Metropolitan and Tier 2 cities preferred, excluding eastern India regions.
3. Age: 35 + ; Gender: Female (Working women)

We applied the above demographic targeting criterion on Facebook audience creation tool to the extent possible. We used higher level education as proxies for high networth individuals and excluded Eastern India regions for targeting, following the nonprofit's suggestion. We detail the specific demographic targeting used on Facebook Audience Creation tool as follows.

Facebook Custom Audience Generation using Firm-provided Demographic Criterion:

- Location: India (*Exclude Eastern India Regions*: Assam, Manipur, Meghalaya, Nagaland, Odisha, Tripura, West Bengal, Sikkim, Arunachal Pradesh, Mizoram, Bihar, and Jharkhand)
- Age: 35 - 65+ ; Gender: Female
- Education Level: Master's degree, Professional degree or Doctorate degree
- Industry: Administrative Services, IT and Technical Services, Legal Services, Sales, Education and Libraries, Business and Finance, Management, Arts, Entertainment, Sports and Media, Architecture and Engineering, Food and Restaurants, Construction and Extraction, Production, Healthcare and Medical Services, Installation and Repair Services, Life, Physical and Social Sciences, Computation and Mathematics, Community and Social Services, Protective Services, Farming, Fishing and Forestry, Cleaning and Maintenance Services, Military (Global), Transportation and Moving or Government Employees (Global)

B. Experiment Details

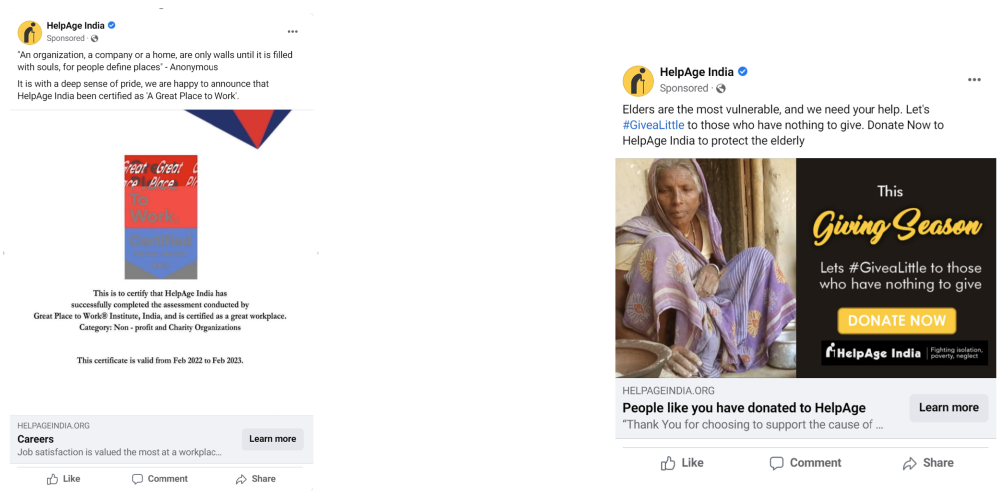
In this section, we provide further information on the randomized experiment setup as follows:

- Ad Placement: Facebook newsfeed (mobile & desktop)
- Objective: Conversion campaign
- Bidding: Dynamic (minimum cost) -Facebook's 'A/B/n' testing feature ensures that each treatment has an equal chance of winning the bid using the lowest cost bid strategy at any given time under a dynamic competition setting. Same ad budget is allocated for each treatment.
- Experiment Duration: Each experiment was conducted for a duration of 7 days. The entire set of experiments were conducted over the period April 2020–March 2021.
- Attribution window: 1-day view, 28-day click

Figure A.1: Sample Ad Design

C. Assessing Ad Lift for Lookalike Targeting

Figure A.2: Ad Creatives for Ad Lift Experiment



(a) Control: Careers Ad **(b)** Treatment

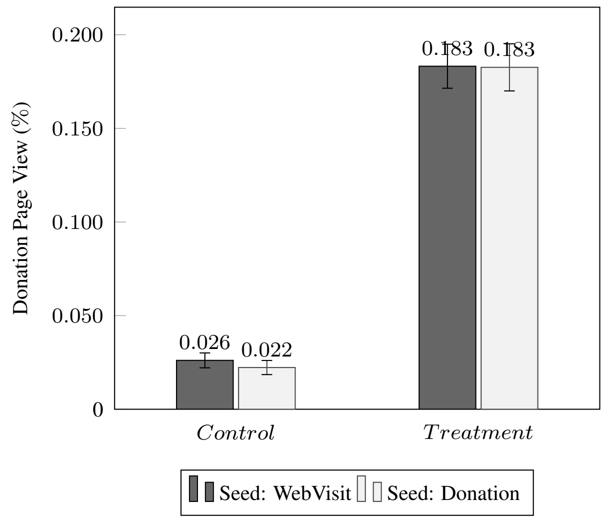
In this experiment, we assess our assumption that without an ad, the baseline visits and donation will be close to zero in the experimental context. We do so with a control condition, where users were directed to a HelpAge careers webpage that listed job openings. In contrast, users allocated to the treatment condition received an ad asking for donations similar to the one used in our main experiment.

The goal of the ad lift experiment is twofold: to show i) that individuals' propensity to visit/donate without solicited advertising is near zero, and ii) the baseline propensity without an ad does not differ between upstream and downstream seeded Lookalike audience.

To the extent that HelpAge careers page is still an ad for HelpAge, there may be some spillover effects on visits and donations to HelpAge, but if we still get close to zero visits and donations with this control, our claim would be valid. Ideally, we would have liked the control to be directed to a public service announcement (PSA) that was not related to HelpAge, but this was not feasible due to Facebook and HelpAge policies.[14]

[14]Facebook does not approve PSA-like holdout advertising from an unrelated business ad account; See https://www.facebook.com/policies/ads/. HelpAge advertising policy does not allow audience traffic to be directed to non-HelpAge content/websites.

Figure A.3: Ad Lift for Lookalike Targeting



Note: Error bars represent standard error.

Figure A.2 displays the ad creatives for the control and treatment conditions. We ran randomized audience and creative split experiments testing effects of advertising for donation (vs. control, i.e., the career ad) with seeding based on Website Visit (upstream) and Website Donation (downstream). The experiment was conducted on a representative sample of 240,832 Facebook users in India.

Figure A.3 shows the results for the control and treatment conditions for the two types of seeding strategy. For both types of seeding, the control condition leads to very small (close to zero) number of visits to the donation page for both website visit seeding (.026%) and website donation seeding (0.022%). Clearly, the treatments effects are much larger at 0.183%. For completeness, we note that donation rates in the control condition were also zero for website visit seeding (0%) and website donation seeding (0.001%). Overall, we conclude that our assumption is valid, and we can interpret the treatment outcomes reported in the paper as very close to actual lift.