Proceedings of the 52nd Hawaii International Conference on System Sciences | 2019

Gauge the Trade-off Effects between Social Media and Traditional Platform in the Consumer Purchasing Funnel

Ran (Alan) Zhang
City University of Hong Kong
Alan.Ran.Zhang@cityu.edu.hk

Daniel Zantedeschi
Ohio State University
zantedeschi.1@osu.edu

Shivendu Shivendu
University of South Florida
shivendu@usf.edu

Abstract

Targeted display advertising for individual consumers has become pervasive on social media platform and other online websites (traditional platform). Yet, the effectiveness of targeted advertising across online platforms is not well understood. Moreover, such advertising effect may be different for different types of consumers, i.e. consumers in the early stage and those in the late stage, relative to the final purchase stage. This paper aims at assessing the effectiveness of targeted advertising across online platforms on consumers' final conversion (purchase). In addition, we measure the complementarity and substitutability of online platforms for targeted advertising for upper funnel (early-stage) consumers and lower funnel (late-stage) consumers. We use machine learning techniques to form case-control designs analyzed employing regularized discrete choice models to select relevant features explaining the final conversion. The empirical analysis shows that (1) targeting across platforms is positively associated with the final conversion for the lower funnel consumers, but there is no measurable synergistic effect for the upper funnel consumers; (2) the main effect of targeting on social media is positively related to the final conversion for consumers in the upper funnel but has no significant impact for lower funnel consumers. We leverage upon these findings to discuss actionable managerial prescriptions.

1. Introduction

The widespread adoption of the Internet and digital technologies has profoundly changed the advertising industry. Within digital advertising spending, display ad spending surpasses search ad spending in the US for the first time [8]. Advertisers invest heavily on display ads that run on various general sites (traditional platform) as well as social media. Social media is an increasingly popular platform and the ad spending on social media is expected to increase from 10.8 billion U.S. dollars in 2015 to 19.3 billion in 2018 [24].

Although advertisers spend a hefty amount of ad budget on social media websites, the effectiveness of display advertising, in particular, targeted advertising on social media is not yet clear to practitioners and academics. On one hand, people naturally connect on social media platform to stay up to date with their social life, e.g. interacting with their families and friends. Thus, they may have little interest in finding advertising useful [25]. On the other hand, social media can provide advertisers with detailed user profile information. The micro-level information becomes a great asset for advertisers allowing them to design and conduct more efficient targeting strategies by displaying customized ads to individual users, leading to potentially higher rates of ultimate conversions [9]. We attempt at answering whether and how targeted advertising on social media can be useful in converting consumers to purchase, relative to that on the traditional platform (Portal website, major media, lifestyle site, etc.). Answering this question will provide insights for academics and practitioners on the effectiveness of targeted ads and help practitioners make an informed decision on effectively allocating their ad budget on different online platforms.

From the advertisers' perspective, it is important to understand whether and how the effects of targeted advertising on social media on consumers' final conversion differ from that on the traditional media. Moreover, it is not clear whether consumers' ad exposure on social media complements with or

URI: https://hdl.handle.net/10125/59524
ISBN: 978-0-9981331-2-6

HICSS

substitutes to that on the traditional platform. It is likely that exposures across the two platforms may have a synergistic effect that is greater than the sum effect of exposures on each platform; Or, it is conceivable that the two platforms may be likely to substitute to each other, so that consumers exposed on both may end up wearing out their interest in the product faster than those exposed on just one platform. In more formal statistical terms, the interaction effect of the two platforms may be different for various types of consumers.

This research is concerned with understanding the effectiveness of targeted advertising on social media relative to that on the traditional platform, and the complementarity or substitutability of platforms for targeted advertising, along with the so-called consumer purchasing funnel (CPF). Specifically, our research questions are in the following:

1. What is the effectiveness of targeted advertising on social media, relative to that on the traditional platform, along with the consumer-purchasing funnel?
2. Does targeting on social media complement with or substitute to targeting on the traditional platform in impacting the final conversions for consumers at different purchasing stages?

Gauging the interaction effects between activities on different platforms and within different parts of the purchasing funnel is very challenging. This is due to (a) the presence of potential activity biases [16], where the most active users end up being targeted more frequently and (b) "rare outcomes" indicating that the ultimate conversion rates are negligible. We tackle these issues by a combination of tools in the epidemiology and machine learning literature comprising (a) case-control design to retrospectively match users presenting a similar level of browsing activities and (b) post-regularized choice models, proved to be effective even in the presence of rare outcomes. We measure the odds ratio to assess the effectiveness of targeting on social platform relative to the traditional platform in both parts of the funnel.

The consumer purchasing funnel is thought to consist of two distinct phases: the upper funnel where users may have some engagement with the firm showing some general awareness of the product, and the lower funnel where consumers have more interaction with the firm showing more interest beyond the general awareness. Consumers can move from upper funnel directly to the purchase stage without going through the lower funnel. Figure 1 exemplifies these two phases showing the presence of different "touchpoints" derived from consumer browsing behavior, as a result of targeted ads; these may happen on either traditional or social media platform.

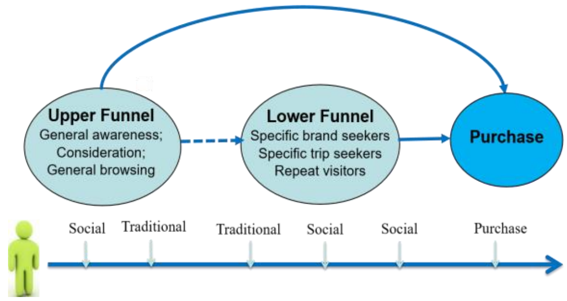

Figure 1: A Consumer's "Journey" to Purchase

From a marketing perspective, customers may be categorized in different purchasing stages depending on the prior history and interaction with the firm. It is important to note that consumers at either funnel stage may be likely to purchase or drop out without buying.

This paper investigate tools that can be employed in the increasing online advertising ecosystem, Specifically, this study aims at providing answers to managerial questions regarding the effectiveness of targeted advertising on different platforms, whether targeting across platforms is beneficial for advertisers, and if so, to which group of consumers. Answering these questions will help practitioners obtaining more efficient targeting strategies across different platforms for different customers, and thereby allocating their advertising budget more wisely.

2. Literature Review

This research topic is related to several emerging and established areas of research on online advertising.

Multichannel Attribution. The first is related to the general problem of " digital attribution" or how to proportionally split the contribution of each platform and touchpoint in the scenario of an ultimate conversion. Several studies examine the attribution problem based on the funnel framework also adopted in this work [1], [18], [19], [28]. [1] map observed consumer behavior to unobserved consumer purchase funnel and developed a hidden Markov model to measure how the change in the previous stage affects the probability of moving to the next stage and the likelihood of conversion. [18] study the carryover and spillover effects of prior touches through the consumer purchase funnel and measure the incremental contribution of multiple channels to conversions. Our paper is complements this stream of literature in that we study the trade-off effects between platforms on consumers' conversion, namely the complementarity and substitutability of targeting on social media and traditional platform on consumers' final purchase.

The effectiveness of Display Advertising. This study is also concerned with the measurement of the effectiveness of display advertising [11], [20]. [20] develop a survival model cast in a hierarchical Bayesian framework to measure the impact of banner advertising on consumers' probabilities of repurchase. [11] employ a smart identification strategy based on a natural experiment, in the context of display ads, and demonstrate that more exposure to display advertising can increase users' propensity to search. Although our study belongs to this general stream, none of the prior research focuses on measuring the effectiveness of targeting on multi-platforms on consumers' final conversion and examine the trade-off effects between platforms on different types of consumers.

The effectiveness of Retargeting Strategies. Another related stream of literature is on the effectiveness of re-targeting. Prior research examines how the effectiveness of re-targeting is affected by information specificity [14], timing and contextual factors [3], and restricting intrusive privacy information [2].

The Complementarity and Substitutability of Channels. This less explored body of research relates to the literature on trade-offs across different channels, i.e., the complementarity or cannibalization effects of digital and physical media. For example, [27] study how offering digital content cannibalizes demand of print circulation. [10] examine the impact of the introduction of digital medium on consumer welfare. [17] study the impact of e-books sales on changes in market coverage and find total market expands when the publisher offers e-books together with print books. Our research contributes to this stream of literature by examining the complementarity and substitutability of social media and traditional platforms in the context of targeted advertising, and we measure the interaction effects of targeting on the two platforms in nonlinear marketing response models.

3. Data and methodology description

The data analyzed in this study is provided by a large international travel & tourism company offering an expensive experience product. The advertiser on behalf of the travel firm runs multiple campaigns and ads for the product on an assortment of websites classified to two platforms, i.e. social media and the traditional platform. Social media includes websites like Facebook and YouTube while traditional platform comprises websites such as Yahoo and AOL. Each ad is associated with one unique campaign and one specific targeting strategy. Every time a user browses a website that belongs to the firm's advertising network, a cookie embedded in the website places a unique identifier in the user's browser. The cookie then tracks the user's viewing and clicking on ads across all websites within the firm's advertising networks. If the user visits the firm's website or makes a purchase, the information is also recorded.

Our individual-level data consists of time information of a user's ad impression (exposure to an ad), clicks (if any), visits of the company's website, and purchases over a period of slightly less than two months. For each creative (ad), we have information about targeting strategy, platform type, ad network, and type of publishers. Since the travel package is an expensive and highly-considered product, we have very small number of purchasing users. Our dataset consists of over 19 million users whose information about their touches were recorded, i.e. type of targeting, on which platform, type of ad networks, etc. Among these users, we only have 1555 consumers who made a purchase: this provides an effective conversion ratio of less than 0.01 percent per cookie chain albeit in line with industry standards. The rarity of the ultimate conversion is the first methodological challenge that we need to address.

3.1. Retrospective Case-Control Methods

To deal with what is commonly called in the epidemiological literature as "rare outcomes" (i.e. the 0.01% effective conversion rate), we develop a retrospectively matched case-control study (see [22] Chapter 4 and 8 for an introduction and comprehensive taxonomy of case-control methods). In the context under consideration, the outcome of interest is "purchases", and targeting on social media and across platforms are the risk factors to be assessed. Compared to the propensity score matching method, the Case-control method is well suited to investigate rare outcomes as it allows for the identification of multiple risk factors associated with these rare outcomes [28].

Since we are interested in examining how targeting on social media and across platforms impact final purchase, our case-control study is retrospective, meaning that given the outcome status, purchase or non-purchase, we "look back" and assess the history of a consumer's online exposures and examine the impact of targeting on social media and across platforms. More specifically, while looking back for each purchase event (cases) we find "similar" set of customers (controls) that ended up not purchasing. The matching procedure is based on information that compares cases and controls by associational variables related to their behavior along the funnel. This is different from the cross-sectional studies, pervasive in IS literature and known as "prevalence" studies. These studies evaluate subjects at one point in time and do not have an inherent temporal dimension.

Following the approach described above, we randomly selected 100,000 non-purchased consumers. These are potential "matches" for our cases/rare outcomes. To find the "match" between case and control groups, we use robust unsupervised learning techniques to identify "similar" consumers who did not purchase, based on the characteristics of those who purchased. Based on the rare outcome hypothesis and properly executed and matched control group, we will be able to obtain (1) estimates that are statistical testable and will preserve the direction of the results within each cluster of consumers detected by the algorithm, and (2) to compare the relative odds ratios between clusters as a measure of prima facie evidence of advertising effectiveness.

An important methodological consideration that we ought to clarify is whether the random subsampling adopted in the first stage of the case-control procedure may end up biasing the numerical magnitude of the estimated coefficients. If the subsample is sufficiently large, sampling biases may not necessarily happen, but in a second order, due to the properties of the odds ratios, it can be shown that the significance and direction of the results within each cluster and the relative odds ratios across clusters are preserved under such subsampling. In the interest of brevity, we refer to [29] that obtained asymptotic results for case-control studies.

3.2. Descriptive Statistics

We present summary statistics in Table 1 and 2 for consumers in the upper funnel and the lower funnel, respectively. The advertiser labels the funnel stage for each consumer at a given time applying the proprietary algorithm to identify the funnel stage based on a consumer's prior browsing history on an assortment of websites. The table 1 and 2 show statistics for each targeting strategy (behavioral, contextual, geo, looklike, predictive, prospecting, retargeting), each platform (traditional, social media), the total time of touches (time length), the inter-time between impressions (inter-time), and the platform for the first touch (fTraditional, fSocial). In general, we find that there is a lot of heterogeneity within the current data but also some distinctive patterns. Interestingly, for consumers labelled as in "the upper funnel", the traditional platform is the most used for targeted ads, while social media platform is most frequently targeted platform for consumers in the lower funnel. Also, while consumers in the upper funnel on average are less exposed to retargeting ads than to behavioral or contextual targeting ads, consumers in the lower funnel receive more retargeting ads than any other targeting type. The websites a consumer visits belongs to one of the advertiser's ad networks (network 1 – 4). Each network comprises a set of websites and the networks differ in the types of purchase contracts the advertiser has with publishers.

Table 1: Descriptive Statistics for Consumers in the Upper Funnel

	Sum	Quan_.5	Quan_.95	Mean	St.dev	Max
No. of Touches	206109	2	12	4.089	5.057	87
Behavioral	69489	0	5	1.379	2.768	52
Contextual	31735	0	3	0.630	1.442	33
Geo	8283	0	0	0.164	1.763	49
Looklike	0	0	0	0.000	0.000	0
Predictive	0	0	0	0.000	0.000	0
Prospecting	47338	0	4	0.939	2.649	67
Retargeting	20500	0	2	0.407	2.554	58
Traditional	136012	2	8	2.698	4.018	86
Social	41576	0	4	0.825	2.628	58
Time Length	224076.1	0.755	19.557	4.446	7.564	61.213
Inter-time	68565.69	0.212	6.213	1.360	2.404	38.317
fTraditional	38129	1	1	0.756	0.429	1
fSocial	11566	0	1	0.229	0.420	1

Table 2: Descriptive Statistics for Consumers in the Lower Funnel

	Sum	Quan_.5	Quan_.95	Mean	St.dev	Max
No. of Touches	352913	3	20	5.705	6.990	86
Behavioral	91978	0	7	1.487	3.973	49
Contextual	32236	0	3	0.521	2.257	49
Geo	84771	0	7	1.370	3.739	50
Looklike	203	0	0	0.003	0.075	4
Predictive	529	0	0	0.009	0.135	9
Prospecting	2996	0	0	0.048	0.386	15
Retargeting	118665	0	11	1.918	5.354	86
Traditional	101536	0	7	1.641	3.600	50
Social	183634	0	15	2.968	6.344	86
Time Length	241564	0.190	18.599	3.905	6.906	60.849
Inter-time	62712.9	0.053	5.016	1.014	2.506	46.122
fTraditional	25153	0	1	0.407	0.491	1
fSocial	26151	0	1	0.423	0.494	1

4. Preliminary Analysis

To further motivate the need for more sophisticated methods to deal with the statistical and managerial problems of consumers targeted in the purchasing funnel, we have performed some preliminary analysis. Specifically, to explore the associations between possible predictors and the probability of purchasing, we perform a *Kitchen Sink Logistic Regression* including all predictors as covariates. These include each type of targeting, platforms, the number of touches, etc.

Table 3: Kitchen sink logistic regression for all predictors ***p<0.001, **p<0.01, *p<0.05

	Estimate	SE	tStat
Intercept	-4.709***	0.430	-10.943
No. of Touches	-0.058***	0.008	-7.139
Contextual	0.664***	0.115	5.746
Geo	-0.610***	0.178	-3.419
Lookalike	1.497***	0.340	4.399
Predictive	1.624***	0.346	4.686
Prospecting	-0.927***	0.109	-8.531
Retargeting	1.375***	0.111	12.370
Traditional Platform	1.777***	0.416	4.270
Across platforms	-2.403***	0.463	-5.188
Social Media Platform	1.760***	0.428	4.116
Time of online path	0.012	0.007	1.804
Inter time btw touches	-0.084***	0.022	-3.863
Path start time	-0.006**	0.002	-2.959
Network 1 (N1)	-0.857***	0.276	-6.870
Network 2 (N2)	-1.459***	0.337	-4.313
Network3 (N3)	-1.831***	0.324	-5.783
Network 4 (N4)	-1.206***	0.190	-6.793

The results in the table show, unsurprisingly in “big data” environments, almost all predictors are significant. A quick inspection of the design matrix can easily reveal collinearity among predictors, i.e., targeting types and platforms; this is because advertisers are most likely to run platform-specific targeting strategies. This hints at the problem of the endogenous targeting assignment to the user likely to create "activity biases" as described in [16]. The following section presents our methodology to address collinearity and activity biases issue based on the case-control method described earlier in combination with the machine learning tools.

5. Methodology

Given that we have a very small number of consumers who purchased (rare events) in our observational data, as discussed earlier, we adopt a retrospectively matched case-control method to measure the odds ratio, equivalent to the relative risk, to assess the effectiveness of targeting on social platform relative to the traditional platform in both parts of the funnel.

In our dataset, we have a very small number of consumers who have moved from upper to lower funnel, and even fewer consumers among them have made purchases. The vast majority of consumers are in either upper or lower funnel in our data time window, thus we focus on these consumers in this study. Because consumers in the upper funnel may have quite different characteristics and prior history from those of consumers in the lower funnel [18], we differentiate upper funnel consumers and lower funnel consumers and form them into two separate groups. This categorization was performed by the advertiser’s proprietary algorithm that labels each consumer based on the consumer’s profile and prior history, prior to the observation window. Note that in each group a small portion of consumers has made purchases.

To find the “matched” control group, we use the robust K-means clustering technique to retrospectively identify “similar” consumers who did not purchase, based on the characteristics of those who purchased. This matching approach enables us to identify the effects of targeted advertising on and across platforms on the final conversion (purchase) of consumers. We use consumer-initiated actions as similarity measures to form clusters of consumers in the upper funnel and in the lower funnel, respectively. Consumer-initiated actions include ad networks that a consumer has visited. A consumer certainly knows which website she is currently visiting, but may not know which ad-network the website belongs to. Compared to the propensity score matching that treats all dimensions equally, the proposed matching procedure selects dimensions with sufficient variation under an orthogonality constraint with other non-selected dimensions such as the targeting strategies.

To identify important predictors and address the collinearity issues between different targeting strategies and platforms, we advocate for regularization methods, particularly the *Elastic Net* logistic regression. After selecting the predictors, we then use the selected covariates and perform a post-regularized logistic model to produce consistent estimates of the odds ratios (See [4] for a general overview of post-regularization methods).

6. Empirical Analysis

We establish the different numbers of clusters for the upper and lower funnel, based on the average similarity between cases and controls in the same cluster. Hence, we experimented with a different number of clusters using elbow method, and eventually, we obtained an optimized number of clusters for each funnel, i.e. 2 clusters for the upper funnel, and 4 clusters for the lower funnel. Each cluster has a different number of users and characteristics of users. Hence, the *Elastic Net* Logistic Model may select a different set of predictors that have more weights of importance than other predictors in predicting the odds ratio of purchase. We report the post-regularized logit model and results for each these clusters. We then present an integrative analysis quantifying trade-off effects between the different platforms.

6.1. Clusters for the upper funnel

For cluster 1 in the upper funnel, the derived post-regularized model is given by:

$$logit(purch_i) = \beta_0 + \beta_1 * Retarg_i + \beta_2 * Social_i + \beta_3 * Length_i + \beta_4 * N1_i + \beta_5 * N3_i + \beta_6 * N4_i + \mu_i \quad (1)$$

where $\mu_i \sim GEV(0,1,0)$.

For cluster one, we can see that relative to targeting on the baseline, traditional platform, targeting on social media has additional significant and positive association with consumers' final conversion (i.e. the odds ratio of purchase). Targeting across the two platforms, however, has no incremental impact relative to targeting on the traditional platform on consumers' final conversion, and the interaction term has been dropped by the regularization. Also, comparing to other targeting strategies (Behavioral, contextual, etc.), retargeting is positively associated with consumers' purchase. Finally, as we expected, the total length of time during ad exposures is negatively associated with the ultimate conversion.

Table 4a: Results of estimation for the upper funnel cluster 1

Upper Funnel, Cluster 1			
	Estimate	**SE**	**tStat**
(Intercept)	-3.3487	0.0876	-38.2480
Retarg.	**1.2284**	**0.2614**	**4.6999**
Social	**0.8935**	**0.1519**	**5.8804**
Length	-0.0443	0.0086	-5.1781
N1	-2.1311	0.2948	-7.2287
N3	-1.5797	0.7137	-2.2134
N4	-1.4018	0.1997	-7.0209

For cluster 2 in the upper funnel, the derived post-regularized model is given by:

$$logit(purch_i) = \beta_0 + \beta_1 * NumTouch_i + \beta_2 * Trad * Social_i + \beta_3 * Social_i + \beta_4 * Length_i + \beta_5 * N1_i + \beta_6 * N3_i + \mu_i \quad (2)$$

where $\mu_i \sim GEV(0,1,0)$.

For cluster two, similar to the result for cluster one, we have that targeting on social media may have additional significant and positive impact on consumers' final conversion, relative to targeting on the traditional platform. On the other hand, targeting across the two platforms has no additional significant effect for consumers in cluster two. Also, we don't observe a significant and positive effect of retargeting strategy. The exposure length of time is negatively associated with consumers' final conversion.

In summary, results for consumers at the upper funnel show that the main effect of targeting on social media has additional significant and positive effect on consumers' conversion, relative to the baseline, targeting on the traditional platform. However, there is no evidence of synergistic effects of targeting across the two platforms. These results suggest that for upper funnel consumers, targeting on social media may have a positive impact on the final conversion, but too much-personalized targeting across platforms may bring no additional impact, perhaps due to that consumers have little willingness to consider purchasing at this stage or have not been familiar with the brand or product. Strategies like retargeting appear to work only for a subset of customers in the upper funnel. Lastly, consistent with our expectation and intuition, the longer the experience in the upper funnel the less likely are customers to ultimately convert.

Table 4b: Results of estimation for the upper funnel cluster 2

Upper Funnel, Cluster 2			
(Intercept)	-3.821	0.4292	-8.9032
NumTouch	-0.17	0.0436	-3.8956
Trad*Social	-0.033	0.4475	-0.0734
Social	**1.176**	**0.428**	**2.75**
Length	-0.216	0.0334	-6.4723
N1	-1.4	0.1552	-9.0161

N3	-0.871	0.4091	-2.1292

6.2. Clusters for the lower funnel

Consumers in the lower funnel are more experienced with the product or have more prior interaction with the firm.

Interestingly, targeting on social media alone is not significant in the lower funnel, suggesting that as consumers have moved to the lower funnel, they may become more sophisticated and might actively search for the product, so targeting on social media alone may not be helpful for moving these consumers to final purchase. However, targeting across platforms is positively associated with the odds ratio of purchase, suggesting that more targeting across platforms may be helpful in providing personalized information and in converting consumers in the lower funnel. Retargeting is significant and positively associated with the odds ratio of purchases in three clusters, suggesting that comparing to other targeting strategies, retargeting appears to be more effective to the lower funnel consumers in helping them convert to the final purchasing stage.

Different from the results for the upper funnel, for consumers at the lower funnel, the average inter-time between two ad exposures is significantly and positively associated with the ultimate conversion, while the number of total ad exposures has a significant and negative association with the final conversion. This may suggest that too frequent exposures may have a negative impact on converting consumers but giving consumers more time to accumulate interest and familiarity with the product or brand may be helpful with the conversion.

It is also interesting to notice the presence of an “empty” set of predictors (risk set): This means that for the people in that group, it was not possible to measure any significant marketing activities determining their purchases consistent with traditional customer base analyses and probabilistic response models (see [26]).

Table 5: Results of estimation for the lower funnel clusters

Lower Funnel			
Cluster 1			
	Estimate	**SE**	**tStat**
(Intercept)	-4.9486	0.4263	-11.6080
NumTouch	-0.1642	0.0614	-2.6750
Retarg.	**5.1171**	**0.5242**	**9.7623**
Trad*Social	**1.7896**	**0.3813**	**4.6931**
Social	-0.4313	0.5366	-0.8038
Length	-0.6866	0.0551	-12.4630
InterTime	0.6581	0.0898	7.3273
N1	1.3421	0.6794	1.9755
N2	-0.1422	0.3166	-0.4492
Cluster 2			
(Intercept)	-5.577	0.548	-10.177
NumTouch	-0.347	0.0932	-3.724
Retargeting	**2.644**	**0.304**	**8.696**
Trad*Social	**3.208**	**0.546**	**5.872**
Social	-0.029	0.4934	-0.0582
Length	-0.898	0.2081	-4.3132
InterTime	0.8502	0.2295	3.705
N1	0.4792	0.587	0.8164
N3	-2.8	0.4048	-6.9174
N4	-1.678	0.3313	-5.0654
Cluster 3			
(Intercept)	-2.79	0.6387	-4.3687
NumTouch	-0.075	0.0578	-1.2962
Retarg.	**2.892**	**0.534**	**5.417**
Trad*Social	0.0638	0.4992	0.1278
Social	-0.406	0.5127	-0.7928
Length	-0.128	0.0164	-7.8106
InterTime	0.0424	0.0294	1.443
N1	-0.722	0.7707	-0.9365
N2	0.1465	0.5279	0.2776
N3	-1.621	0.6685	-2.4244
N4	-0.841	0.5477	-1.5362
Cluster 4			
(Intercept)	-4.5	0.159	-28.304

7. Tradeoff measures between social media and the traditional platform

In the context of targeted advertising, consumers visit different sites and thereby may be exposed to targeted ads on both social media and the traditional platform at different times before making a purchasing decision. For example, a consumer might first receive targeted ads on the traditional platform, and then receive targeted ads on social media, and get targeted again later on the traditional platform. As is known in the literature, there is an interaction effect of two independent variables on the dependent variable, if the effects of the two independent variables are more (or less) than the sum of the parts. The interaction of the independent variables also underlies moderation effects [7]. In our context, the interaction effect of social media and the traditional platform implies that the effect of targeting on social media on the log odds ratio of purchase is moderated by the effect of targeting on the traditional platform, and vice versa. Estimation of the interaction term is also at the center of our analysis as we wish to understand whether, ex-post, the multichannel targeting strategies delivered by the agency were effective in delivering ultimate conversions.

Interpreting moderation effects in nonlinear models are often not straightforward. To examine the relationship between targeting on social media and the traditional platform, however, we do need to interpret the interaction effects in a more qualitative and insightful manner (see [7], also mentioned in the previous section). Our logistic regression targets the "relative risk" (RR) framework for assessing the importance of different risk factors from well-established epidemiological literature.

It is well known in the statistical literature that approximations for interaction analysis exist under the rare outcomes assumption [23]. This allows us to estimate interaction effects and interpret interactions in a "linear" probability scale and leverage about the notion of relative risk described above. In particular, we consider the Relative Excessive Risk due to Interaction (RERI)[1]. We calculate RERI based on the following:

$$RERI = RR_{11} - RR_{10} - RR_{01} + 1$$

Subscript "11" refers to activating targeting on both social media and the traditional platform, "10" refers to putting social media but shutting down the traditional platform, and "01" refers to shutting down social media while activating the traditional platform. RERI is presented in a more familiar linear and additive form (thus avoiding the cumbersome problem of inverting log-odds) and can be interpreted qualitatively as the "extra lift" of the probability of purchase due to the presence of the ads on both platforms. Specifically, If $RERI > 0$, social media and traditional platform are considered complement; If $RERI < 0$, social media and traditional platform are considered a substitute.

Table 6: RERI table, *p < 0.05

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Upper Funnel	-0.01	0	N/A	N/A
Lower Funnel	0.04*	0.05*	0.02*	0

It is easy to notice that social media and the traditional platform is the lack of synergistic effects in the upper funnel and more as a complement for consumers in the lower funnel. These results point out the possibility of the complex complementarity patterns that could be better exploited by the firm when delivering the ads.

8. Conclusion

We have developed an empirical strategy with the aim of identifying interaction effects between activities performed on different platforms within different parts of the funnel. First, our results indicate that targeting across platforms has synergistic effects with the ultimate conversion for consumers at the lower funnel, but does not appear to provide any interaction effect for the upper funnel consumers. Second, our results show that the main effect of targeting on social media, relative to that on the traditional platform, is positively associated with the odds ratio of purchase for the upper funnel consumers, but has no significant relative impact for consumers at the lower funnel. Lastly, our findings indicate that the commonly implemented "retargeting ads" are more effective than other more sophisticated targeting strategies, and that retargeting may have a positive and significant association with the ultimate conversions for consumers at the lower funnel.

Finally, our study draws managerial implications by measuring the trade-off effects between social media platform and the traditional platform for digital advertising. Our findings help answer the important managerial questions regarding what platform(s) the advertiser should run their advertisement on, through what targeting strategy, and for which type of consumers. Specifically, targeting on social media may be more helpful and can bring incremental informational value when consumers are at early

[1] See [30]. We also note that we could call the "RERI" as the "interaction lift."

stages. However, we do not detect a synergistic effect for targeting across platforms when consumers are not experienced or familiar with the products or brand. We speculate that too much-personalized ads across platform may not be helpful, or bring negative psychological impact on early stage consumers, even though the informational value is positive. When customers move to a more mature purchasing stage, targeting across platforms appears to be very beneficial. Finally, based on our findings and suggestions, advertisers may consider allocating more ad budget on retargeting than on other targeting strategies such as behavioral or contextual targeting, which often involves complicated negotiations and implementations across different ad networks, and it may be more efficient to retarget consumers who are at the lower purchasing funnel.

This work can be extended in several aspects. First, we have ignored consumers who experienced both funnels due to the small number of consumers with purchases. It may be interesting to include this group of consumers in the future study to examine how targeting on different platforms affects the probability of consumers transitioning in purchasing funnels. Second, we may need to characterize selected clusters in a more policy interpretable manner. Third, we could potentially extend our approach to account for the effects of the mobile platform on conversions, and to measure the associational effects of different types of ads across platforms.

Acknowledgement

This research received data from Wharton Customer Analytics Initiative (WCAI) and funded by Marketing Science Institute (MSI).

References

[1] Abhishek, V., Fader, P., & Hosanagar, K. (2012). Media exposure through the funnel: A model of multi-stage attribution. *Available at SSRN 2158421*.

[2] Aziz, A., & Telang, R. (2016). What is a digital cookie worth?. *Available at SSRN 2757325*.

[3] Bleier, A., & Eisenbeiss, M. (2015). Personalized online advertising effectiveness: the interplay of what, when, and where. *Marketing Science*, 34(5), 669-688.

[4] Chernozhukov, V., Hansen, C., & Spindler, M. (2015). Valid post-selection and post-regularization inference: An elementary, general approach. *Annual Review of Economics*, *7*(1), 649-688.

[5] Chintagunta, P. K., & Haldar, S. (1998). Investigating Purchase Timing Behavior in Two Related Product Categories, *Journal of Marketing Research*, 35, 43-53.

[6] Dellarocas, C., Gao, G., & Narayan, R. (2010). Are consumers more likely to contribute online reviews for hit or niche products?. *Journal of Management Information Systems*, 27(2), 127-158.

[7] Dhar, T., & Weinberg, C. B. (2015). Measurement of interactions in non-linear marketing models: The effect of critics' ratings and consumer sentiment on movie demand. *International Journal of Research in Marketing*.

[8] eMarketer (2016). U.S. Digital Display Ad Spending to Surpass Search Ad Spending in 2016. http://www.emarketer.com/Article/US-Digital-Display-Ad-Spending-Surpass-Search-Ad-Spending-2016/1013442

[9] Ganguly, S. (2015). Why Social Media Advertising is set to Explode in the Next Three Years? http://marketingland.com/social-media-advertising-set-explode-next-3-years-121691

[10] Gentzkow, M. (2007). Valuing New Goods in a Model with Complementarity: Online Newspapers. The *American Economic Review*, 713-744.

[11] Ghose, A., & Todri, V. (2015). Towards a Digital Attribution Model: Measuring the Impact of Display Advertising on Online Consumer Behavior. *Available at SSRN (2015).*

[12] Greenland, S. & Thomas, D.C. (1982). On the Need For the Rare Disease Assumption in Case-Control Studies. *American Journal of Epidemiology* 116(3): 547-553.

[13] King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political analysis*, 9(2), 137-163.

[14] Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research*, 50(5), 561-576.

[15] LePage, E. (2015). A Long List of Social Media Advertising Stats. https://blog.hootsuite.com/social-media-advertising-stats/

[16] Lewis, R.A, Rao, J.M & Reiley, D.H. (2011). Here, there and everywhere: correlated online behaviors can lead to overestimates of the effects of advertising. *Proceedings of the 20th international conference on world wide web.* Pages 157-166.

[17] Li, H. (2015). Cannibalization or Market Expansion? The Impact of E-Books on Print Book Sales. Working paper. Carnegie Mellon University. Pittsburgh, PA United States.

[18] Li, H., & Kannan, P. K. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, 51(1), 40-56.

[19] Li, Y., Xie, Y., & Zheng, E. (2015). Competitive Analytics of Multi-channel Advertising and Consumer Inertia. *Working paper*. https://pdfs.semanticscholar.org/7995/ec143d20711ec7108a0e3cae86bda3459c6b.pdf

[20] Manchanda, P., Dubé, J. P., Goh, K. Y., & Chintagunta, P. K. (2006). The effect of banner advertising on internet purchasing. *Journal of Marketing Research*, 43(1), 98-108.

[21] Manski, C. F. (1999). Nonparametric Identification Under Response-Based Sampling, in C. Hsiao, K. Morimune, and J. Powell, eds., Nonlinear Statistical Inference, Cambridge University Press.

[22] Rothman, K. J., Greenland, S. and Lash, T. L. (2008). *Modern Epidemiology*, 3rd Edition, Wolters Kluwer Editions.

[23] Richardson, D. B., & Kaufman, J. S. (2009). Estimation of the relative excess risk due to interaction and associated confidence bounds. *American journal of epidemiology*, kwn411.

[24] Statista (2018). Social network advertising spending in the United States from 2015 to 2018. https://www.statista.com/statistics/736971/social-media-ad-spend-usa/

[25] Sibley, A. (2015). 11 Examples of Facebook Ads that Actually Work (and Why). http://blog.hubspot.com/blog/tabid/6307/bid/33319/10-Examples-of-Facebook-Ads-That-Actually-Work-And-Why.aspx

[26] Schmittlein, D. C., Morrison, D. G., & Colombo, R. (1987). Counting your customers: Who-are they and what will they do next?. *Management science*, 33(1), 1-24.

[27] Simon, D. H., & Kadiyali, V. (2007). The effect of a magazine's free digital content on its print circulation: Cannibalization or complementarity? *Information Economics and Policy*, 19(3), 344-361.

[28] Song, J. W., & Chung, K. C. (2010). Observational studies: cohort and case-control studies. *Plastic and reconstructive surgery*, 126(6), 2234.

[29] Xie, Y., & Manski, C. F. (1989). The logit model and response-based samples. *Sociological Methods & Research*, 17(3), 283-302.

[30] Vanderweele, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*, Oxford University Press.