Effects of Internet Display Advertising in the Purchase Funnel:

Model-Based Insights from a Randomized Field Experiment

Paul R. Hoban[1]

Randolph E. Bucklin[2]

February 2015

ABSTRACT

We study the effects of internet display advertising using cookie-level data from a field experiment at a financial tools provider. The experiment randomized assignment of cookies to treatment (firm ads) and control conditions (charity ads) enabling us to handle different sources of selection bias, including targeting algorithms and browsing behavior. We analyze display ad effects for users at different stages of the company's purchase funnel (e.g., non-visitor, visitor, authenticated user, converted customer). We find that display advertising positively affects visitation to the firm's website for users in most stages of the purchase funnel, but not for those who previously visited the site without creating an account. Using a binary logit model, we calculate marginal effects and elasticities by funnel stage and analyze the potential value of reallocating display ad impressions across users at different stages. Expected visits increase almost 10 percent when display ad impressions are partially reallocated from non-visitors and visitors to authenticated users. We also show that results based on the controlled experiment data differ significantly from those computed using standard correlational approaches.

Keywords: Internet advertising, field experiment, purchase funnel, logit model

---

[1] Assistant Professor of Marketing, Wisconsin School of Business
975 University Ave
Madison, WI 53706
e-mail: phoban@bus.wisc.edu
phone: 517.410.6188

[2] Professor of Marketing, Peter W. Mullin Chair in Management, UCLA Anderson School
110 Westwood Plaza
Los Angeles, CA 90095, USA
e-mail: randy.bucklin@anderson.ucla.edu
phone: 310.825.7339

Each quarter in the United States, more than one trillion internet display ads are delivered and nearly 300 individual advertisers spend at least $1 million on them (comScore 2013; comScore 2011). A $17.7 billion dollar industry in 2013, spending on display advertising is expected to increase 53 percent by 2017, to $33.4 billion per year (eMarketer 2013).

Despite this growth, managers often question whether online display advertising truly affects customer behavior. General Motors recently terminated a $10 million ad spend with Facebook while publicly questioning whether such ads could influence consumer behavior (Terlep et al. 2012). Wine.com CEO Rich Bergsund moved his company's entire display ad spend to paid search, affiliate marketing, and comparison shopping engines. His concern was that display ad impressions, "may not convert to customers or sales," (Barr and Gupta 2012). These concerns highlight the need to develop and test reliable methods to establish and measure the links between online display advertising and consumer behavior. This is critical not only for advertisers, but also for the internet companies who depend upon display ad revenue and the online platforms which manage the sale and distribution of display advertising.

One challenge in studying the effects of online display ads is that selection effects are widespread – along with the corresponding risks of bias in estimated effect sizes. Correlating exposure to display ads with behavior outcomes can be problematic because the users exposed to the ads (versus those who are not) may have had those ads targeted to them based on their propensity to act in certain ways. For example, if users interested in financial services are more likely to browse sites with financial content, ads for financial services firms placed alongside such content or targeted to individuals regularly consuming such content may spuriously appear to be more productive. User browsing behavior can also make exposure-based correlates problematic because users who visit more sites and view more pages may be more likely to both see an ad and visit the web site.

Intertwined with these selection issues is the increasingly common practice of retargeting display ad impressions (i.e., disproportionately focusing ad impressions on prior visitors to a website). Firms offering such services point to significant increases in ad effectiveness, attributed to reducing shopping cart abandonment and re-engaging past customers (AdRoll.com 2014). However, recent research has shown that the effect of increased targeting may not always be positive, and can depend on the product category (Goldfarb and Tucker 2011). Further, the effectiveness of different appeal types varies with consumer browsing behavior and product preferences (Lambrecht and Tucker 2013). Together, these findings indicate that firms also need to consider how display advertising effectiveness may change with the consumer's relationship to *their* product and firm – i.e., stage of the purchase funnel. It is not clear at what stage display ads are likely to be most effective for a firm. Are they better at bringing new users to a company's website (i.e., an awareness building role), encouraging those already familiar with the site to come back (i.e., retargeting), or reminding users with an established relationship or previous purchase to return or buy again?

The purpose of this paper is to extend our understanding of internet display advertising by analyzing its effects by stage in the purchase funnel. Following previous researchers (e.g., Goldfarb and Tucker 2011), we use field experiment data to address the problem of selection bias. We harness detailed cookie-level data collected during the experiment to track purchase funnel stage by user along with display ad exposure. In our application these stages are non-visitor, visitor, authenticated user, and converted customer but our approach can readily accommodate others. Our goal is to shed light on the potential for display ad effects to differ by funnel stage while also illustrating a practical approach to analyzing this aspect of display ad response. We also provide implications for the allocation of ad impressions to purchase funnel stage.

In what follows, we present model-free and model-based analyses of data from a large scale, randomized field experiment conducted by a collaborating firm. This firm sells online financial management tools directly to consumers. Our data include tracking records of individual consumer exposure to display advertising along with those individuals' browsing behaviors at the firm's website, information that is identical to what is generally available to managers. In the experiment, a small proportion of individuals were randomly assigned to the control group. This group was served ads in precisely the same manner as the treatment group, but the ad copy was for an unrelated charity instead of the focal firm. For the control group, any correlation between display advertising and the outcomes of interest cannot stem from advertising effects, but would be due to potential selection bias from confounds such as individual level targeting and browsing behavior. By comparing the ad response between the treatment and control groups, we estimate the effect of display advertising while controlling for selection bias. Importantly, we are able to do this while reserving only 1.6% of impressions for the control group, an important cost consideration in the use of randomized field experiments and for managers looking to implement our approach.

We begin our analysis of the experimental data by presenting model-free evidence. We find that the impact of display ads varies by purchase funnel stage. It is positive and significant for three of the four stages but it is ineffective for users who have previously visited the firm's website and did not create an account. Though these users were exposed to detailed product information during one or more prior site visits, they declined to proceed. This may indicate that the (relatively) low information content of the display ads was insufficient to persuade these individuals to reconsider their decision not to move further down the funnel.

Using an individual-level advertising response model, we estimate marginal effects and elasticities while incorporating lag effects and holding constant factors such as timing and seasonality. Comparing our results based on the experiment with those obtained from a simple correlational approach, we find significant differences in estimated display ad effectiveness. We also examine the optimal ad impression allocation across funnel stages that would be implied by our elasticity estimates and find that they differ dramatically from the allocation actually employed by the firm. Our results suggest that reallocating a significant proportion of the ad exposures from individuals in early funnel stages to those who have created an account but not yet purchased is expected to increase site visits by nearly 10 percent.

BACKGROUND AND LITERATURE

The need for a clear link between online display advertising and a firm's outcomes of interest has not gone unnoticed in the literature. In some of the earliest work on display advertising, Chatterjee et al. (2003) focused on understanding what drove click through rates. Using data from an online content provider, they found that new visitors and less frequent visitors to the site showed a stronger propensity to click. At about this time, click through rates dropped precipitously, and researchers sought the underlying causes and better metrics of banner ad effectiveness. Dreze and Hussherr (2003) showed that although individuals actively avoid looking at display ads, they still have a positive effect on brand awareness and advertising recall. Cho and Cheon (2004) established perceived goal impediment, the belief that the ad is not relevant to the objective at hand, as the underlying cause of display ad avoidance. This was supported by Danaher and Mullarkey (2003) who found that banner ads have more influence on individuals who are browsing than on those who are performing a goal directed activity.

Rutz and Bucklin (2012) examined and quantified the link between display ad exposure and brand interest. Using data from a third party automotive site, they showed that display ad exposure can influence within site browsing behavior. Specifically, consumers were significantly more likely to seek content on the site related to previously advertised brands than those that were not advertised. In line with previous work, they found that effect sizes varied by browsing behavior, with users who created less focused click streams showing a greater response to advertising.

Manchanda et al. (2006) moved beyond the classical brand-based measures of ad effectiveness, directly linking banner ad exposure and purchase behavior. Using a semi-parametric hazard model, they found that banner ad exposure had a positive effect on purchase frequency for existing customers. This effect was greatest when consumers viewed a large number of web pages across a variety of websites (i.e., when across site browsing increased).

Lewis et al. (2011) pointed out that heavy browsers are also more likely to perform a wide variety of online behaviors independent of advertising exposure. Because the probability of ad exposure increases with browsing duration and intensity, unaccounted for correlation between browsing behavior and the outcomes of interest will create a confound (Cameron and Trivedi 2005). In their application, Lewis et al. found that correlational measures would have led to significant overestimation of display advertising's influence on these behaviors of interest. The results from Lewis et al. highlight the need to use caution when interpreting findings about display advertising that may not have controlled for selection bias.

In addition to browsing behavior, potential selection bias also stems from the individually targeted nature of much display advertising. Since the early days of online advertising, managers and researchers have sought to increase ad effectiveness by targeting individuals with relevant browsing and search histories (Sherman and Deighton 2001). In recent years, major ad servers have begun aggregating user histories across services, including email, search, and social networking (Ingram 2012). Because the targeting algorithms used by ad servers are frequently only partially observable to firms (Google 2012), accounting for such targeting is complex. For instance, firms may request that a certain ad be retargeted to prior site visitors, but they may not know how the ad server selects which prior visitors will receive impressions.

While one might assume that targeting would, at worst, have no effect on individual responses, Goldfarb and Tucker (2011) showed that targeting can negatively affect display ad effectiveness when paired with highly visible creative. Using data from a large scale field experiment, they found that display ads that were both obtrusive and targeted had less impact on purchase intent than those that were only one or the other. They showed that this effect is most pronounced in categories generally considered private (financial products, healthcare, etc.) and for individuals who seem to most value privacy.

Lambrecht and Tucker (2013) also examined the impact of targeted display advertising, focusing on the dynamic relationship between appeal type effectiveness and consumer preferences. Using a quasi-experimental design and data from an online travel provider, they found that dynamically retargeted ads (i.e., those that show specific products the consumer previously viewed) generally underperform more generic appeals. However, this relationship is reversed when a consumer has visited an online review site, which is interpreted as a refinement in product preferences. This indicates that there exists a dynamic relationship between display advertising effectiveness and the consumer’s position in the purchase funnel. An objective of our study is to further explore ad effect differences by funnel stage by taking advantage of data from a randomized field experiment where multiple stages are observed.

In sum, two selection issues may bias standard correlational estimates of display advertising effectiveness. First, individuals who browse more web pages are, *ceteris paribus*, more likely to see a given ad and may be more likely to perform a number of activities, such as visiting a site and buying online. Second, individual level targeting may lead to individuals with higher baseline probabilities of site visits being served more impressions. In either case, there would be a positive bias in effect size estimates. The bias, however, need not always be positive. There may exist activities, such as making critical life decisions, that are less likely to be undertaken when users are actively browsing a wide variety of web pages. Similarly, recent work has shown that certain types of advertisements perform worse when heavily targeted, with a pronounced effect in the financial and health arenas. Consequently, the two sources of selection bias, browsing behavior and targeting, can have opposing effects with unknown magnitude. Thus, it is unclear *a-priori* whether correlational measures will overstate or understate the effect of display advertising in any given application.

APPROACH

Our proposed approach is based upon the conduct of a controlled, randomized field experiment (so-called "A/B testing") and the model-free and model-based analysis of individual, cookie-level tracking data from it. Our data come from a large scale experiment in which a small proportion of individuals were randomly assigned to the control group. This group is targeted in precisely the same manner as the treatment group, but is shown ad copy for an unrelated charity in place of firm advertising. Throughout the experiment, ad impressions and site visits are observed at the individual level using tracking cookies. This allows us to identify a baseline independent of the firm's message and copy, while still following the same ad serving (and therefore targeting) algorithms. Through our combination of random assignment, identical targeting, and individual level tracking, we are able to control for selection bias.[1]

As an alternative to a controlled experiment, advanced correlational approaches might be used to handle endogeneity due to targeting and browsing. Instrumental variables and control functions use additional covariates to create an orthogonal relationship between the regressors and the error term (Cameron and Trivedi 2005). For targeting, this would require both knowledge of the proprietary targeting algorithms used by ad servers and access to the relevant covariates, neither of which are generally available. In the past, researchers have attempted to control for browsing behavior bias by including the number and variety of other websites an individual is observed to visit (Manchanda et al. 2006). However, these counts are only observable when a firm's ad is served during the site visit. This can be an effective control in untargeted campaigns, as was the case for Manchanda et. al. (2006). However, the correlation between this measure of browsing behavior and any unaccounted for targeting can introduce an additional source of endogeneity.

In dealing with missing variables and individual level targeting, our problem mirrors that of Manchanda et al. (2004), who studied pharmaceutical detailing. They simultaneously estimate targeting and response models while allowing the individual level response parameter to enter the targeting algorithm. Unfortunately, their approach is not a good fit for our application. To handle the unobserved nature of browsing behavior, we would need some exogenous predictor. As mentioned above, such predictors are generally unavailable to the researcher. Second, identification of the parameters in their model requires significant individual level variation in both targeting and response. While this may be possible for very large campaigns, we found the model to be empirically unidentified in our application.[2]

The experiment was set up as follows. During their first digital interaction with the firm (either through ad exposure or site-visit), individuals were randomly assigned to either the treatment or control group. When determining when and to whom a display ad impression should be served, the ad servers did not differentiate between the two groups and the targeting was consistent between them. When selecting the ad copy to be served, the control group was always given copy unrelated to the firm; in our case, these were ads for a large multi-national charity. Thus, any apparent effect of display advertising in the control group represents the combined effect of selection biases such as from targeting or browsing behavior. For the treatment group, the apparent effect of display advertising is the joint effect of these factors plus the actual effect of online display advertising. Thus, the effect of display ads on behavior is the *difference* in the probability of the action between members of the treatment and control groups. While field experiments with randomized treatment and control conditions can be costly (due to the opportunity cost of foregone impressions served to users in the control group), we are able to gauge effect sizes by purchase funnel stage from an experiment with a modest 1.6 percent of cookie ids assigned to the control group.

From the within site browsing behavior, we identify four key stages in the purchase funnel (non-visitor, visitor, authenticated user, and converted customer), allowing us to directly examine how advertising's impact may change as individuals move through the purchase process. As discussed above, Lambrecht and Tucker (2013) found that the effectiveness of specific versus generic retargeted ads shifted with consumer preferences. Previous research also supports a dynamic relationship, finding that brand familiarity increases advertising impact and delays copy wear-out (Campbell and Keller 2003; Kent and Allen 1994), and product experience dominates advertising effects (Hoch and Ha 1986). However, for many products and services (i.e., cars, bank accounts, etc.), consumers may gain both brand familiarity and product experience as they move through the purchase funnel, making it unclear *a-priori* whether subsequent advertising will be more or less effective. Our approach allows for these effects to be examined empirically.

As we will show, the model-free evidence of ad effectiveness will be illuminating as it is broken out across stages of the purchase funnel. Nonetheless, this analysis alone does not fully unlock the value of the online experimental data. We augment the model-free analysis by developing a binary logit response model, estimated in a Bayesian framework. By controlling for the number of impressions, modeling the experimental data allows us to accomplish several additional objectives without the need to increase the size of the control group.

Specifically, the model allows us to control for additional covariates and incorporate lags and impression counts. Using these counts, we can also calculate marginal effects and elasticities for display ad exposure. Leveraging this information, we then evaluate our approach versus correlational approaches. Finally, we use our effect size estimates to show how impressions could be reallocated across stages of the purchase funnel so as to improve response outcomes at the same cost. In sum, our experimentally-based approach accounts for selection biases stemming from multiple sources, reveals ad effectiveness at different stages of the purchase funnel, and can be easily implemented by firms at a relatively low cost.

DATA

The data for this study were provided by a large financial services firm, and focus on a single suite of consumer financial management tools. From the customer’s perspective, this product line is largely independent, maintaining its own brand and website. Due to the nature of our research agreement with the company, we cannot share the name of the firm or the precise details of their business. The data are at the individual cookie level, and detail the online interactions between the firm and individuals in a mid-size Midwestern market during the six week period from February 19 to April 2, 2010.

Because our data come from a single firm advertising one product during a specific time in a given market, we do not draw general conclusions regarding display advertising effects. Instead, we use our data to show how the approach can be applied to reveal new findings about display ad effects and how the results might be used to improve advertising productivity. We leave the further generalization of our empirical findings as a topic for future research.

Data based on tracking cookies also have a number of known limitations. A tracking cookie is tied to a unique browser. It is possible that multiple users share a browser; in this case our data are analogous to household level panels. An individual also may use multiple browsers, resulting in multiple cookies. Toupet et. al. (2012) compare cookie data from a common analytics provider to the proprietary Nielsen panel; they find that the vast majority of individuals are associated with only a single cookie. Also of concern is that individuals who allow tracking cookies to persist on their machine may differ in some critical, unknown respect from those who disable or frequently delete them. As pointed out by Chatterjee et. al. (2003), many websites block access to browsers with cookies disabled, leaving consumers with little practical choice in this regard. With respect to cookie deletion, research has also shown that most individuals delete cookies less than once per month (comScore 2007). Further, Dreze and Zufryden (1998) used a randomized experiment to show that consumer browsing behavior is not significantly influenced by the presence of tracking cookies.

In our data, each display impression contained one of nine different designs, consisting of the same solid color background with a phrase, icon, or both in the foreground. Unfortunately, we do not have sufficient information to consistently identify the creative for each impression. While we have been assured by management that the only systematic variation in creative was by week (which we will control for), this precludes us from pursuing questions related to the efficacy of various designs or appeal types. The ads were distributed across myriad sites, using targeting algorithms that were only partially known to the firm. For example, the firm could control the total number of impressions per day or retarget individuals that had previously visited the site, but they were not privy to all of the individual level factors determining when and to whom impressions were served.

Based on the web pages a user visits within the firm's site, we identify four stages in the purchase funnel: non-visitor, visitor, authenticated user, and converted customer. Similar to many ecommerce websites, a consumer must proceed sequentially through these stages to complete an online transaction with the firm, though they may move through any number of them during a single visit. Non-visitors are individuals who have never interacted with the firm's website. Visitors have been to the site, but have not provided the personally identifiable information necessary to sign-up for an account. Authenticated users have signed up for an account, but no money has changed hands. Finally, a converted customer has completed a transaction.

We have varying expectations about how consumers in each funnel stage may respond to display advertising for the firm. For example, those who have never been to the site (non-visitors) may be less familiar with the brand, and thus ads may help build awareness but have small effect sizes. In contrast, those who have been to the site but who have not progressed to establishing an account may have decided that they will not do so. In such cases, display advertising may have little or no effect. Authenticated users (those who have signed up but not purchased) are likely to be very aware of the brand, and the ads may serve as a reminder to complete a transaction. We do not have a clear expectation about the effect on converted customers, but because of the interest among some consumers in follow-up, we conjecture that the ads may also serve as a reminder at this stage. Managerial considerations also motivated our choice of purchase funnel structure, because the firm also uses these stages to collect click stream metrics and to study the path to conversion.

We use site visit as our dependent variable of interest. We define this as any time a user visits the firm's website, regardless of the actual within-site browsing behavior. Our data come from a large provider of web-based financial management tools, and these tools are deeply integrated into the firm's website so as to provide a uniform look and feel. This means that consumers experience many of the product attributes (ease of use, clarity of messaging, etc.) immediately upon site visit. Prior research has shown that such direct product experience dominates advertising's impact on attitudes, beliefs, and behavior (Hoch and Ha 1986; Marks and Kamins 1988; Tellis 1988; Wright and Lynch 1995). Because users begin evaluating product attributes at the point of site visit, we believe that display advertising's influence on subsequent actions such as sign-up or conversion will be overshadowed.[3] In addition, using site visit as the dependent variable allows us to compare display advertising's impact at various funnel stages, where we find significant, informative differences. Finally, site visit allows for some generalizability of our approach across websites. While we use site visit as the dependent variable in this study, it is important to note that the approach could easily be repurposed to measure the effect of display ads on other observable click stream behaviors. We discuss one such alternative, advancement to the next funnel stage, in the *Model-Based Results* section below.

During our observation period, we track 133,058 cookies, 2,164 (1.6%) of which were randomly assigned to the control group. Because the data contain only observed interactions between the user and the firm, we only know that a cookie is actively tracking a user between the first and last observed events. Thus, including lag effects and funnel position effects can create an initial conditions problem. To mitigate these issues, we require an ad exposure or site visit at least three weeks prior to our focal period and use this initialization period to determine the funnel position. We also require that each cookie survive for at least four weeks, because frequent cookie deletions would bias our estimated effects towards zero.[4] We culled cookies in the top one percent of total impression counts to remove crawlers (machines programmed to download large numbers of web pages without human interaction). We also removed individuals who were not served a display ad impression, because they cannot be identified in our dataset as charity or firm. Finally, we removed observations from 795 mobile devices.

The data were discretized to the cookie-day, resulting in 4,748,020 observations. During this period, 2,216,947 display ad impressions were served, 33,096 (1.5%) of which were served to our control group. Table 1 presents the distribution of observed impression counts, showing a significant mass at zero and a long tail. The vast majority of cookie days (87.5%) contain no impressions, and 98.2% of observations contain five or fewer impressions. Table 2 contains the observation counts by treatment group and funnel stage. Note that the treatment group has a larger proportion of observations in later funnel stages. This follows expectations if advertising has a positive impact.

(TABLE 1 ABOUT HERE)

(TABLE 2 ABOUT HERE)

MODEL-FREE RESULTS

To obtain model-free results, we examine the difference in probability of site visit between the treatment and control groups for each funnel stage. We aggregate the cookie-day observations by treatment group and funnel stage, then calculate the probability of site visit within each group, where site visit is a binary indicator for each day. (Note that we assume that multiple page views on a single day are a part of the same visit.) The results are presented in Figure 1. In each stage, cookies in our treatment group are more likely to visit than those in the control group. Non-visitors, visitors, authenticated users, and converted customers in the treatment group are respectively 0.07%, 0.01%, 0.99%, and 0.52% more likely to visit on any given day than their control group counterparts. While these differences may seem small, this is due to the low overall probability of site visit. In terms of odds, this translates to a respective 74.7%, 0.6%, 49.7%, and 48.2% increase between the control and treatment groups. This difference is significant ($p \leq 0.05$) for non-visitors and authenticated users and marginally significant ($p < 0.10$) for converted customers.

(FIGURE 1 ABOUT HERE)

These results provide strong model-free evidence for the efficacy of online display advertising and differences in that efficacy along the purchase funnel. Nevertheless, there are a number of limitations. First, they do not directly control for potentially important observable factors, such as lagged effects or timing. Second, they do not account for the number of impressions served to users. Third, they do not provide insight into marginal effects or elasticities, which are needed for making allocation decisions.

These limitations are not a result of the experimental approach per se – all of these effects could be estimated non-parametrically given a sufficiently large sample. Rather, they stem from practical constraints associated with the control group cost, which would increase exponentially as experimental factors are added. In a balanced experiment, the required sample size generally increases exponentially with the number of additional factors to avoid a loss in precision. Given the disparate size of our treatment and control groups, we may be able to obtain equal precision by increasing only the control group. Unfortunately, the incremental cost of our approach is directly proportional to the size of the control group. Measuring marginal effects further complicates matters, because each impression count would represent an additional treatment level and require a sufficiently large sample size to reliably estimate effects. Unfortunately, firms cannot dictate individual impression counts. Thus, no finite sample can guarantee reliable non-parametric estimation of marginal effects. With statistically significant results for only two of four groups, we appear to be pushing the bounds of what can be learned non-parametrically, even with 4,748,020 observations and 2,216,947 display ad impressions.

MODEL

To control for potentially important factors, estimate lag effects, and calculate marginal effects and elasticities, we turn to an individual-level model of advertising response. Given the binary nature of our dependent variable, whether or not a site visit occurred, we turn to a standard binary logit estimated at the cookie-day level. Below, we will discuss how we use this model to account for selection biases as well as the limitations created by the sparse nature of our data.[5]

Using the standard binary logit model, we express the probability of site visit at time t conditional on a set of covariates X as:

$$Pr(visit_t|X,\beta) = \frac{exp(X\beta)}{1 + exp(X\beta)}. \tag{1}$$

Here, X contains the log transformed contemporaneous and lagged impression counts for both firm and charity ads, indicators for funnel stage, date controls, and interactions between the logged contemporaneous effects and funnel stage indicators.[6] β is the vector of parameters to be estimated. The parameters for contemporaneous and lagged impression counts as well as the related interactions differ depending upon whether the user is in the treatment or control group. We specify the remaining parameters to be equal between groups because group assignment is random and, therefore, the only difference is in the ad shown. This specification allows us to maximize the information content on which the ad response parameters are estimated, conditional on a given control group size. We log transform the impression counts to incorporate decreasing marginal effects of advertising and to reduce the influence of large impression counts on parameter estimates. This specification also provides a better fit than either linear or squared transformations, as measured by the Bayesian Information Criterion (BIC).

As we will discuss in more detail below, our estimated effect size is the difference between the fitted probability of site visit for members of the treatment and control group. While a point estimate for any set of covariate values could certainly be calculated using a frequentist approach, the Bayesian approach more readily provides us with the posterior distributions necessary to evaluate the estimated effect sizes. Further, it allows us to easily aggregate posterior distributions over groups of covariate values. In this way, we can look at the effect of display advertising by funnel stage, impression count, or both.

To obtain actual model-based effect sizes, we must deal with a series of issues. First, the effect of an interaction term is not equal to its marginal effect in non-linear models, but rather the magnitude, direction, and significance are all a function of the remaining covariate and parameter values (Ai and Norton 2003). Second, the effect of n exposures to a firm's advertising is the probability of site visit at time t conditional on those exposures minus the probability of site visit for an otherwise identical individual conditional on the same number of exposures to control ad impressions. This difference in probabilities is not easily interpreted from any one of the marginal posterior distributions.

Because the Bayesian approach produces draws from the joint posterior distribution of the parameters, we can overcome these challenges. Using the joint posterior draws, we can simulate the predictive distribution for any set of covariate values, $\Pr\left(visit_t \mid x_{t,f}, x_{t,c}, \beta, \Psi \right)$.[7] Here, $x_{t,f}$ is the number of firm impressions at time t, $x_{t,c}$ is the number of control impressions, Ψ represents the vector containing the value of all other covariates, and β represents a matrix containing all draws from the joint posterior distribution of the parameters in Equation (1). As discussed above, we are not directly interested in these values, but rather the change in probability of site visit, v, given n firm impressions relative to the probability of site visit given the same number of charity impressions. Let $\Delta_v(n)$ represent this difference in predictive distributions such that,

$$\Delta_v(n) = Pr\left(visit_t \middle| x_{t,f} = n, x_{t,c} = 0, \beta, \Psi\right) - Pr\left(visit_t \middle| x_{t,f} = 0, x_{t,c} = n, \beta, \Psi\right). \tag{2}$$

To include K lags, we generalize Equation (2) by calculating the difference in probability of a site visit in any of the K periods given n firm impressions less the probability of a site visit given n control impressions. This gives the following:

$$\begin{aligned} \Delta_v(n) = & \bigcup_{k=0}^{K} \Pr\left(visit_{t+k} \middle| x_{t,f} = n, x_{t,c} = 0, \beta, \Psi\right) \\ & - \bigcup_{k=0}^{K} \Pr\left(visit_{t+k} \middle| x_{t,f} = 0, x_{t,c} = n, \beta, \Psi\right). \end{aligned} \tag{3}$$

Note that we take the union of the probabilities over the K periods instead of the sum to avoid double counting the probability of a site visit occurring on multiple days.[8]

Given this effect size estimate, we can simulate the predictive distributions for the marginal effects and elasticities. For n impressions, the incremental effect of an additional impression, $\mathcal{M}_v(n)$, is the effect of $n+1$ impressions minus the effect of n impressions:

$$\mathcal{M}_v(n) = \Delta_v(n+1) - \Delta_v(n). \tag{4}$$

Following the standard definition, advertising elasticity can be calculated as the percentage change in advertising effectiveness divided by the percentage change in impression count. Because impression counts are by nature an integer value, the elasticity of an additional impression reduces to the product of $\mathcal{M}_v(n)$, n, and the inverse probability of site visit given n firm impressions:

$$\eta_{x_{t,f}} = \Delta_v(n) \times \left(\frac{n}{\bigcup_{k=0}^{K} Pr\left(visit_{t+k} \middle| x_{t,f} = n, x_{t,c} = 0, \beta, \Psi\right)} \right). \tag{5}$$

Following equations (2) through (5), we can simulate the predictive distributions for a given set of covariate values. However, in evaluating the size and robustness of an effect, it is often more informative to look at these distributions over a range of covariate values. To do this, let Ψ be a matrix containing the fully enumerated values over the observed span of X. With some abuse of notation, equations (3) through (5) then produce a matrix of draws from the relevant predictive distributions, where each row represents the distribution for a given set of covariate values.[9] We use these joint predictive distributions in what follows.

MODEL-BASED RESULTS

We specified and estimated four versions of the model. Analogous to the model-free calculations, each specification was estimated at the cookie-day level, with site visit as the outcome of interest. Our estimates are based on 100,000 draws, with a 10,000 iteration burn-in period, retaining every 10th. Our first model, Model 1, contains only the information available in the model free results:

indicators for funnel stage, group assignment, and their interactions. The results (presented in Table 3) are similar to what we see in the model free evidence, with null effects for visitors only.

(TABLE 3 ABOUT HERE)

(TABLE 4 ABOUT HERE)

Model 2 begins to unlock the full value of the experimental data by controlling for logged contemporaneous firm and charity impression counts, purchase funnel stage, and the associated interactions. The results for this model (and all subsequent specifications) are provided in Table 4. As mentioned above, directly interpreting the parameter estimates for interacted terms in Table 4 can lead to incorrect conclusions regarding the significance, direction, and magnitude of their effects. However, the table does give two key pieces of information. First, we can interpret the estimated effects of non-interacted terms within this model (e.g., the large negative intercept estimate (-6.649) is a result of the low average frequency of site visitation). Second, the final row contains the Bayesian Information Criterion (BIC), from which we can clearly see that controlling for contemporaneous impression effects significantly improves model fit (193,547 for Model 2 versus 197,445 for Model 1). In fact, this 3,898 point drop represents 81 percent of the total decline between Model 1 and Model 4 (our preferred specification), meaning that most of the power in our model stems from including this impression count information.

In Model 3, we allow the model to capture spurious correlation between ad response and timing by including weekly indicators, as well as an indicator for weekend versus weekday. Weekdays during the sixth week serve as our baseline, so the significant parameter estimates for weeks one (0.590), two (0.471), three (0.165), and five (0.170) indicate that subjects were generally more likely to visit during these weeks than during the sixth week, *ceteris paribus*. Similarly, the positive parameter estimate for weekend (0.108) shows that subjects were more likely to visit on Saturday or Sunday than the remaining days, *ceteris paribus,* likely fitting with the time demands (and privacy needs) related to using an online financial management tool. We also see that controlling for the timing of impressions in this way provides a significantly better model fit, with a BIC 798 points lower than Model 2.

Finally, Model 4 allows prior day impressions to influence site visit. We see that these effects are negative and significant for both firm impressions (-0.159) and charity impressions (-0.501), possibly indicating some negative correlation in the unobserved browsing behavior between days. That the effect is less negative for firm than charity impressions indicates that the display ad partially mitigates some of this negative correlation. With a drop of 144 points in BIC, Model 4 provides a significantly better fit than Model 3.

We tested a number of additional model specifications, focusing on extending the lag structure and allowing for interactions between contemporaneous funnel position and lagged impressions. Given the sparse nature of the data and the disparate size of the control and treatment groups, we must be careful in evaluating model fit. As shown in Equation (2), the effect of online display advertising is the difference in the fitted probability of site visit between otherwise identical members of the treatment and control groups. Model parameters must then be added in pairs (one each for the treatment and control groups) to identify an effect. However, standard model selection methods evaluate only the overall model fit, and remain agnostic as to which parameters drove the improvement. Given the relatively small control group, additional parameters may over fit the data. This may result in wide confidence bands for the probabilities of site visit, masking effect sizes. To account for this, we tested model extensions by adding control group parameters first and continuing with treatment group parameters only if model fit improved as measured by BIC. We found that neither an expanded lag structure nor interactions between contemporaneous funnel position and lagged impressions improved model fit.

To examine how display advertising's effect varies by purchase funnel stage, we fully enumerate the span of our data, including up to six focal and lagged impressions.[10] Because we are modeling the effect of display advertising conditional on funnel stage, we require funnel stage to remain constant within an observation. That is, an individual in our simulated data cannot be an authenticated user during the first period and a converted customer in the next. If such a transition were allowed, our estimates of display advertising effectiveness would be confounded with the main effects for funnel stage. This restriction is largely consistent with our data, as only 0.7% of observed impressions carry over from one funnel stage to another. We also restrict non-focal impressions (i.e., lagged impressions at time t and contemporaneous impressions at time $t+1$) to be firm impressions. Without this constraint, the effect of focal impressions would be confounded with that of non-focal impressions. Thus, we are measuring the impact of additional firm impressions given some consistent state. Using this simulated data, our posterior draws, and equations (3) through (5), we can plot the posterior distributions for the effects, marginal effects, and elasticities of display advertising for any set of covariates.

Figure 2 contains the posterior predictive distribution of effect size for a single display advertising impression, broken down by funnel stage. Similar to Figure 1, we see that the effect of display advertising is positive and significant for non-visitors, authenticated users, and converted customers, while it has no discernible effect on prior visitors. The null effect for this last group may be attributable to the dominant influence of product experience over advertising effects.[11] In visiting the site, these individuals have been exposed to detailed firm offerings, but declined to proceed. It is unlikely that the firm's display advertising, a monochrome banner containing a single icon, tagline, or both, will contain sufficient new information to persuade these individuals to reconsider. Comparing the distributions across funnel stages, it is clear that the effect size is significantly smaller for non-visitors than for authenticated users or converted customers, with at least 99.9% of the posterior distributions for these later stages being greater than the median draw for non-visitors.[12] Given the randomized nature of the field experiment, these differences provide additional evidence that consumer response to advertising changes with familiarity and experience.

(FIGURE 2 ABOUT HERE)

Alternative Outcome Measures

Directly relating display advertising to a particular desired final outcome (e.g., sales revenue or profits) can be difficult, if not infeasible due to the very low rates of sales conversion and other measurement challenges. In these situations, intermediate measures of advertising effects need to be used. While we use site visit as the dependent variable in this study, our approach can be repurposed to measure the effect of display ads on other behaviors recorded in the click stream. Where possible, studying multiple intermediate outcomes can provide some assurance that the findings are not driven by specific limitations of a particular dependent variable. With this in mind, we estimated an additional model using transition to the next funnel stage as an alternate dependent variable: for non-visitors, this is visiting the website; for visitors, this is creating an account; and for authenticated users, this is completing a purchase. For converted customers, there is no clear interpretation, so they are dropped from this analysis.

Figure 3 reports the effects sizes for the model with next funnel stage transition as the dependent variable. The estimated effect sizes are similar to those produced using site visit as the dependent outcome of interest. With 98.8% of the posterior predictive distribution greater than zero, display advertising has a significant, positive impact on the probability of purchase among authenticated users. With a median estimated effect of 0.48%, the effect is just less than 1/3 of the 1.47% increase in the probability of site visit for this group estimated in Model 4. Display advertising also has a positive, significant impact on the probability of site visit among non-visitors, with 98.2% of the posterior predictive distribution greater than zero. At the median, we estimate that a single display ad increases the probability of site visit by 0.15% for non-visitors, which is roughly equivalent to the 0.11% we estimate in Model 4. The effect is again insignificant for visitors, though it is now directionally positive at the median. These similarities provide reassurance that our approach and findings are robust to the selected outcome of interest.

(FIGURE 3 ABOUT HERE)

Marginal Effects & Elasticities

Figures 4 and 5 contain the median and 95% confidence bands for the effect and marginal effect at each of the purchase funnel stages for up to six impressions. Each point on the plot represents the increase in the probability of site visit from one additional impression. All of the previous findings are supported, and we now see a marked difference in marginal effects between visitors and both authenticated users and converted customers. While the effect of a display ad impression is small for non-visitors, so is the rate of decreasing marginal effectiveness. For authenticated users and converted customers, the effect sizes are larger, but degrade more quickly with additional impressions.[13]

(FIGURES 4-6 ABOUT HERE)

Figure 6 plots the posterior distribution of the elasticity estimates at a single impression. The descriptive statistics for these distributions can be found below in Table 5. While the display advertising effect is smaller for non-visitors than converted customers, the median elasticity estimates are almost equal. Note that by examining the elasticity at a single impression, impression count drops out of Equation (5), resulting in elasticity being equal to the ratio of the marginal effect and the probability of site visit given a single firm impression:

$$\eta_{1_{t,f}} = \frac{\Delta_v(1)}{\bigcup_{k=0}^{K} Pr\left(visit_{t+k} \middle| x_{t,f} = 1, x_{t,c} = 0, \beta, \Psi\right)}. \tag{6}$$

Since the marginal effect for non-visitors is significantly smaller than for converted customers, the comparable elasticity estimates are due to the relatively small probability of site visit for non-visitors.

(TABLE 5 ABOUT HERE)

Excluding visitors, all posterior distributions for elasticities are positive and significant, with at least 99.9% of their posterior draws greater than zero. For the three funnel stages with significant elasticity estimates, the combined 95% confidence band spans 0.05 to 0.16 with a median of 0.10. Interestingly, this is exactly in line with the average sales-to-advertising elasticity of 0.10 found in existing literature (Hanssens 2009). However, we caution that these are not sales-to-advertising elasticities, but rather site-visit-to-advertising elasticities. Further, this estimate focuses on the three responsive segments, ignoring prior visitors. The closest comparison to our measure is likely that of Rutz and Bucklin (2012), who report a page-choice-to-on-site-advertising elasticity of 0.20 for within site browsing behavior. Notably, their estimate focuses on browsing behavior within the same website, while ours focuses on driving behavior across sites.

Examining First Impression Stage Observations

In our experimental setting, the random assignment of cookies takes place at the first digital interaction between user and firm. This means that users who have progressed from one funnel stage to another remain in the same experimental condition. A concern with this approach is that treatment group users in later funnel stages may have been exposed to firm ad impressions while their control group counterparts have not. To investigate whether this affects our findings, we proceed as follows. First, we examine the subset of observations in which the user's funnel stage is the same as when they are served their initial ad impression. These so-called "first impression stage" observations enable us to ensure that the effect of display advertising is isolated from exposures at earlier funnel stages. Comparing the site visit probabilities between treatment and control groups, we find no significant differences from Figure 1. We also conducted a model-based assessment using interaction terms for first impression stage status and found no meaningful differences in our results.[14] Thus, we conclude that our findings are robust to this issue. We attribute this robustness to the very low carryover effects for display ads. Nonetheless, we encourage investigators to conduct similar checks of their own data. When carryover effects are stronger, the experiment can be modified by randomly reassigning group membership at each stage of the purchase funnel.[15]

Optimal Allocation & Expected Site Visits

Based on the elasticity estimates, we can calculate the optimal allocation and estimate the resulting expected number of visits using the Dorfman-Steiner condition. This condition states that, at optimality, marginal revenue must equal marginal cost for each marketing instrument (Dorfman and Steiner 1954; Lambin et al. 1975; Sridhar et al. 2011). Given this, it follows that the ratio of advertising elasticities for any two marketing instruments is equal to the ratio of their costs. While the Dorfman-Steiner condition is often thought to apply only to monopolies, it holds if we apply the weaker assumption of monopolistic competition (Lambin et al. 1975). That is, we need only assume that the firm's competitors will not react to the recommended shift in advertising allocation. Our recommendations for reallocation are based on a consumer's position in the purchase funnel, as determined by prior within-site browsing behavior. Because this is not observed by the firm's competition, we believe that competitors would be highly unlikely to respond to the recommended shifts in advertising spending.

To be sure, one can imagine scenarios in which a (partially) unintentional competitive response could occur. When a firm reduces the advertising to a given segment (i.e., non-visitor and visitor), these impression opportunities do not cease to exist. Rather, they are reallocated to other advertisers based on the ad server's sales and targeting processes. In a system in which all advertisers sell competitive products, such impression opportunities would certainly be filled by a competitor, and a competitive response would, by definition, occur. However, the vast majority of ad servers work with a wide variety of advertisers, most of whom are almost certainly unrelated to our focal firm. Further, these advertisers are only partially privy to the sales and targeting algorithms used by the ad servers. Thus, the chance of a competitor filling foregone impression opportunities is quite small.

We can apply Dorfman-Steiner using only our existing data, provided we make three additional assumptions. First, we must view online display advertising that is targeted to each purchase funnel stage as a distinct marketing instrument, precluding carryover effects between stages. Given that our preferred model contains a single day lag, only 0.7% of all observed impressions carryover from one stage to the next. Second, we must assume that the cost per impression is constant across funnel stages. This assumption is necessary because management indicated that the firm was not allocating impressions following the Dorfman-Steiner condition and our data does not contain cost information. If they had been targeting impressions optimally or costs were observed, we could relax this assumption. For firms making allocation decisions, costs will generally be observable and should be used in place of impression counts in our approach. Finally, we must assume that the firm's objective is to maximize site visitation, with profit maximization as an assumed byproduct. In our application, consumers begin experiencing the product when they arrive at the site, and, as previously discussed, product experience dominates advertising as a determinant of consumer beliefs and behaviors. Further, there are near zero marginal costs related to each site visit. Therefore, maximizing site visitation is likely closely correlated with maximizing profit, at least for non-converted customers. Because this assumption does not hold for converted customers, we exclude them in the allocation recommendations that follow. Though we have made a number of assumptions to enable this analysis, some of which might be challenged, our intent remains primarily to illustrate how dramatically advertising allocation can shift when effects are understood at the funnel stage level.

Accepting the above, the Dorfman-Steiner condition states that the ratio of site-visit-to-advertising elasticities for any two stages must be equal to the ratio of their total impression counts:

$$\frac{\eta_{nv}}{\eta_{au}} = \frac{Imps_{nv}}{Imps_{au}} \text{ and} \tag{7.a}$$

$$\frac{\eta_{v}}{\eta_{au}} = \frac{Imps_{v}}{Imps_{au}}. \tag{7.b}$$

Where η is the site-visit-to-impression elasticity for a given funnel stage, $Imps$ is the corresponding impression count, nv represents non-visitors, v represents visitors, and au represents authenticated users. Further, we know that the total number of impressions available is the sum of the impressions served to each group, or:

$$IMPS = Imps_{nv} + Imps_{v} + Imps_{au}. \tag{8}$$

Given the elasticity and total available impressions, we solve this system of equations for the optimal allocation. The optimal allocation rules are:

$$Imps_{nv} = \frac{IMPS}{\frac{\eta_v}{\eta_{nv}} + \frac{\eta_{au}}{\eta_{nv}} + 1}, \tag{9.a}$$

$$Imps_{v} = \frac{IMPS}{\frac{\eta_{nv}}{\eta_{v}} + \frac{\eta_{au}}{\eta_{v}} + 1}, \text{ and} \tag{9.b}$$

$$Imps_{au} = \frac{IMPS}{\frac{\eta_v}{\eta_{au}} + \frac{\eta_{nv}}{\eta_{au}} + 1}. \tag{9.c}$$

Because the proposed approach produces a (barely) negative elasticity estimate for prior visitors, we fix this value at zero in what follows.[16]

Table 6 contains the actual impression allocation as well as that based on the median elasticity estimates from the proposed approach. Based on our approach, the firm should dramatically reallocate impressions among funnel stages. All impressions previously served to prior visitors should be allocated to authenticated users, along with 26% of those that were previously served to non-visitors. In the end, the optimal allocation for authenticated users based on our estimates is 497% larger than the actual allocation.

(TABLE 6 ABOUT HERE)

This finding deserves further discussion. The recommended reallocation may be a result of suboptimal allocation on the part of management, or it may be a result of cost differences that are unobservable to the researcher. We cannot disentangle these drivers given our data. While we assume that costs are constant across purchase funnel stages, and indeed it seems unlikely that targeting any given individual would be more expensive than another, this does not fully account for untargeted impressions. Given that the low probability of site visit results in a large proportion of non-visitors relative to other funnel stages, untargeted advertising is more likely to reach non-visitors. If the cost of untargeted impressions is sufficiently small, it may be optimal to allocate more impressions to non-visitors than our approach would indicate. However, this is a result of our assumption of equal costs across funnel stages, and not a result of our estimated effect sizes. For a firm implementing our approach, such cost differences would be observable, and should be included in the optimal allocation calculations.

Using our proposed elasticity estimates provided in Table 5, we can calculate the expected number of visits given these allocation shifts as follows:

$$VISITS = \sum_s E[visits_s] = \sum_s visits_{s,actual}\left[1 + \eta_s\left(\frac{imps_{s,opt}}{imps_{s,act}} - 1\right)\right] \tag{10}$$

where E$[visits_s]$ is the expected number of visits from funnel stage s, $visits_{s,actual}$ is the number of visits observed for that funnel stage in the data set, η_s is the elasticity estimate from our proposed approach for that funnel stage, $imps_{s,opt}$ is the number of impressions served to that funnel stage under the optimal allocation, and $imps_{s,act}$ is the observed number of impressions served to that funnel stage. Table 7 compares the actual number of site visits at each stage, with that expected given our approach. In total, our proposed allocation results in 1,241 expected additional site visits, a 9.85% increase.

Simply increasing expected visits is insufficient to justify our methodology; we must also outperform what would be expected if the 31,758 impressions that were shown to the control group[17] are instead used for firm advertising. These estimates are contained in the last column of Table 7. Given the observed proportion of impressions allocated to each funnel stage, the additional impressions increase expected visits, but only by 16 (0.12%). Given that this is a tiny proportion of the increase delivered by the optimal allocation, the firm is better off following our approach.

(TABLE 7 ABOUT HERE)

Comparison with Correlational Estimates

As noted earlier, selection effects can bias common correlational estimates of display advertising's effectiveness (Goldfarb and Tucker 2011; Lewis et al. 2011). While the combined direction and magnitude of these biases is unknown *a-priori*, we can examine them *a-posteriori*. To obtain comparable standard correlational estimates, we use the same approach as above, except that the simulated control group contains no impressions. This is similar to comparing individuals who saw the firm's advertising to those who did not.[18] In this case Equation 3 becomes:

$$\Delta_v(n) = \bigcup_{k=0}^{K} \Pr\left(visit_{t+k} \middle| x_{t,f} = n, x_{t,c} = 0, \beta, \Psi\right) - \bigcup_{k=0}^{K} \Pr\left(visit_{t+k} \middle| x_{t,f} = 0, x_{t,c} = 0, \beta, \Psi\right). \tag{11}$$

By excluding the effects of advertising on the control group, we give up control for potential biases. Figures 7 through 9 compare the posterior distributions for each measure (effect, marginal effect, and elasticity) from this correlational model to those from our approach. In general, the correlational estimates are biased towards zero for each of the funnel stages, with visitors standing out as a possible exception. For each combination of funnel stage and measure, Table 8 reports the percent of the posterior distribution from our approach which exceeds the median from the correlational approach. For authenticated users and converted customers, the 95% confidence bands from our approach exclude the median from the correlational estimate for each measure. For non-visitors and converted customers, the evidence is less clear.

(FIGURES 7-9 ABOUT HERE)

(TABLE 8 ABOUT HERE)

Even where the correlational estimates are significantly biased, we note that the differences are much smaller than the two to three orders of magnitude reported by Lewis et al. (2011). Also, the direction of bias is reversed. The smaller difference may be due to the outcome of interest we use – site visits. Lewis et. al. selected response measures with a high baseline probability (i.e., page views and searches on Yahoo.com), which are likely to be more highly correlated with browsing behavior than are visits to a lower traffic website. The directional difference may be due to our specific application in financial services. Activities requiring significant concentration and/or privacy, such as making important financial decisions, may be negatively correlated with browsing a large number of web pages.

We can also assess the benefits of our approach versus the correlational one using the Dorfman-Steiner condition. Again, this analysis is intended to be illustrative and we do not observe costs. Table 9 compares the optimal allocations based on the median elasticity estimates from the correlational approach to those in the data and to those recommended by the proposed approach. Under the correlational approach, 373,544 non-visitor impressions are reallocated to later funnel stages, with 23.3% going to visitors and the remaining 76.7% to authenticated users. The recommended impression allocation for authenticated users is 175% larger than the actual allocation but 54% smaller than our approach. Building on Table 7, Table 10 contains the expected number of site visits given this impression allocation and the elasticity estimates from Table 5. While expected site visits increase by 2.78%, this is less than one third the increase delivered under our approach.

(TABLE 9-10 ABOUT HERE)

CONCLUSION

Our purpose was to develop and test an experimentally based approach for estimating the effects of online display advertising that accounts for the selection effects widespread in this domain and enables managers to examine differences in display ad effectiveness by purchase funnel stage. Our dataset for an online financial tools provider contained display ad impressions and firm interactions for 133,058 individual cookies over a six week period. A small subset of these individuals (2,164), were randomly assigned to a control group and shown ads for an unrelated charity in place of the firm's advertising. The control group allows us to handle the selection biases arising from factors such as display ad targeting and heterogeneous browsing behavior.

In our data, model-free results show that online display advertising has a small, positive, and significant effect for three out of the four purchase funnel stages studied. Leveraging variation in when and how many ads were served for additional power, our model-based results provide additional findings, including elasticities, carryover effects, and marginal effects of increasing exposure. In particular, for non-visitors, we find marginal returns to more impression exposures decrease slowly, suggesting an awareness building role for ads at that stage. For authenticated users and converted customers, the marginal effects decay much more quickly, suggesting a reminder role. For individuals who have previously visited the site but declined to provide identifying information, online display advertising has no discernible impact on behavior.

We also examined the implications for impression allocation across funnel stages, showing in an illustrative analysis that substantial increases in site visit would occur if impressions were reallocated based on relative elasticities. Optimal impression allocation across funnel stages based on our proposed approach results in 1,241 additional expected site visits, representing a 9.85% lift in total visits. We also show that the estimated effects derived from our approach and those derived from correlational methods significantly differ.

Our approach depends critically on the use of a randomized field experiment to disentangle display ad effect sizes from selection effects associated with unobserved browsing behavior and individual level targeting. Our Bayesian modeling approach complements this experimental design, allowing us to examine effects that are imperceptible in the model free results due to the sparse nature of internet tracking data. However, those leveraging our method still need to think carefully about the tradeoff between sample size (especially for the control group), the number of desired control variables, and the rarity of the outcome measure. This final point is of particular import for online tracking data, which is generally extremely sparse. For any given cookie-day, there is only 0.3% chance of site visit within our data. Examining less common outcomes, including those further down the purchase funnel such as purchase, would require that an increased proportion of impressions be served to the control group, which comes at a cost. Ultimately, these decisions need to be driven by the perceived purpose of the advertising, and how the information will ultimately be used in future decisions.

As with all experiments, the conclusions are only of value for as long as the underlying data generating process has not changed. In our application, the targeting algorithms, which are unknown and potentially evolving, could be a source of such change. If the targeting methodology were to change during the experimental period, our approach would still yield unbiased estimates of display advertising effectiveness during that time. Further, identifying such a regime switch and estimating its impact is actually one of the strengths of our approach, requiring only a few additional parameters in the model. Specifically, one could interact the week indicators with the logged impression counts. If the change in targeting had a material impact on advertising effectiveness, one would expect two distinct results. First, the extended model would be preferred to more parsimonious counterparts that do not allow for changes in advertising effectiveness over time. Second, impressions served in weeks following a regime shift would be estimated to be more or less effective than their earlier counterparts. In this way, our approach can not only handle changes in targeting methodology that occur during the experimental period, but it can actually be used to identify when they occur.

While we have sought to develop a practical methodology for measuring online display advertising and obtaining managerial insights into its effects, there remain several opportunities to build on our research. First, we do not have information on other marketing variables (e.g., offline advertising, pricing, etc.) during the time of our data. While we would expect such effects to be consistent across the randomly assigned treatment and control groups, such information could serve as a valuable control or provide insight into how effects vary based on observed heterogeneity. Second, we cannot link ad copy to a given exposure due to technical limitations in our data. While we have been assured by management that the only systematic variation was by week (which we control for), this does prevent us from pursuing interesting questions regarding ad copy effects such as how display advertising response in each funnel stage varies by appeal type. Third, we do not have any additional relevant behavior or demographic information. Such information would be helpful in reducing uncertainty surrounding our parameter estimates, and may allow future researchers to disentangle browsing behavior from targeting biases. Finally, our data stem from a single campaign, for a single firm advertising a single product line in one market. Larger, more diverse data sets are needed before our substantive results can be freely generalized or the underlying mechanisms driving differences in display advertising response by funnel stage are fully understood.

REFERENCES

AdRoll.com (2014), "How Retargeting Works," (accessed June 3, 2014), [available at http://www.adroll.com/retargeting].

Ai, C. R. and E. C. Norton (2003), "Interaction Terms in Logit and Probit Models," *Economics Letters*, 80 (1), 123-29.

Barr, Alistair and Poornima Gupta (2012), "Analysis: Microsoft Loss Reflects Web Display Ad World's Woes," *Reuters (*July 8) (accessed July 12, 2012), [available at http://www.reuters.com/article/2012/07/08/us-advertising-internet-idUSBRE86706H20120708].

Cameron, Adrian Colin and P. K. Trivedi (2005), *Microeconometrics : Methods and Applications*. Cambridge ; New York: Cambridge University Press.

Campbell, M. C. and K. L. Keller (2003), "Brand Familiarity and Advertising Repetition Effects," *Journal of Consumer Research*, 30 (2), 292-304.

Chatterjee, Patrali, Donna L. Hoffman, and Thomas P. Novak (2003), "Modeling the Clickstream: Implications for Web-Based Advertising Efforts," *Marketing Science*, 22 (4), 520-41.

Cho, Chang-Hoan and Hongsik John Cheon (2004), "Why Do People Avoid Advertising on the Internet?," *Journal of Advertising*, 33 (4), 89-97.

comScore (2013), "U.S. Digital Future in Focus," (April 2), [available at http://www.comscore.com/Insights/Presentations_and_Whitepapers/2014/2014_US_Digital_Future_in_Focus].

comScore, Inc. (2007), "Cookie-Based Counting Overstates Size of Web Site Audiences," *(*April 16) (accessed January 3, 2012), [available at http://www.comscore.com/Insights/Press_Releases/2007/04/comScore_Cookie_Deletion_Report].

---- (2011), "U.S. Online Display Advertising Market Delivers 1.1 Trillion Impressions in Q1 2011," *(*May 4) (accessed June 13, 2012), [available at http://www.comscore.com/Press_Events/Press_Releases/2011/5/U.S._Online_Display_Advertising_Market_Delivers_1.1_Trillion_Impressions_in_Q1_2011].

Danaher, Peter J. and Guy W. Mullarkey (2003), "Factors Affecting Online Advertising Recall: A Study of Students," *Journal of Advertising Research*, 43 (03), 252-67.

Delo, Cotton (2012), "Microsoft Files Patent to Serve Ads Based on Mood, Body Language," *Ad Age Digital (*June 12, 2012) (accessed June 13, 2012), [available at http://adage.com/article/digital/microsoft-files-patent-ad-serving-tech-senses-mood/235336/].

Dorfman, Robert and Peter O. Steiner (1954), "Optimal Advertising and Optimal Quality," *The American Economic Review*, 44 (5), 826-36.

Dreze, Xavier and Francois-Xavier Hussherr (2003), "Internet Advertising: Is Anybody Watching?," *Journal of Interactive Marketing*, 17 (4), 8-23.

Dreze, Xavier and Fred Zufryden (1998), "Is Internet Advertising Ready for Prime Time?," *Journal of Advertising Research*, 38 (3), 7-18.

eMarketer (2013), "Advertisers Continue Rapid Adoption of Programmatic Buying," *(*November 26) (accessed October 20, 2014), [available at http://www.emarketer.com/Article/Advertisers-Continue-Rapid-Adoption-of-Programmatic-Buying/1010414].

Goldfarb, Avi and Catherine Tucker (2011), "Online Display Advertising: Targeting and Obtrusiveness," *Marketing Science*, 30 (3), 389-404.

Google (2012), "Advertising Policy FAQ," (accessed May 17, 2012), [available at http://www.google.com/policies/privacy/ads/#toc-doubleclick].

Hanssens, Dominque M. (2009), *Empirical Generalizations About Marketing Impact : What We Have Learned from Academic Research*. Cambridge, Mass.: Marketing Science Institute.

Hoch, S. J. and Y. W. Ha (1986), "Consumer Learning - Advertising and the Ambiguity of Product Experience," *Journal of Consumer Research*, 13 (2), 221-33.

Ingram, Matthew (2012), "Should Google's New Privacy Policy Concern You?," *Business Week (*January 25) (accessed June 13, 2012), [available at http://www.businessweek.com/technology/should-googles-new-privacy-policy-concern-you-01252012.html].

Kent, Robert J. and Chris T. Allen (1994), "Competitive Interference Effects in Consumer Memory for Advertising: The Role of Brand Familiarity," *Journal of Marketing*, 58 (3), 97-105.

King, Gary and Langche Zeng (2001), "Logistic Regression in Rare Events Data," *Political Analysis*, 9 (2), 137-63.

Lambin, Jean-Jacques, Philippe A. Naert, and Alain Bultez (1975), "Optimal Marketing Behavior in Oligopoly," *European Economic Review*, 6 (2), 105-28.

Lambrecht, Anja and Catherine Tucker (2013), "When does Retargeting Work? Timing Information Specificity," *Journal of Marketing Research*, 50 (5), 561-76.

Lewis, R.A., J.M. Rao, and D.H. Reiley (2011), "Here, There, and Everywhere: Correlated Online Behaviors Can Lead to Overestimates of The Effects of Advertising," in Conference Proceedings from International World Wide Web Conference: Hyderabad, India.

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.

Manchanda, Puneet, Peter E. Rossi, and Pradeep K. Chintagunta (2004), "Response Modeling with Nonrandom Marketing-Mix Variables," *Journal of Marketing Research*, 41 (4), 467-78.

Marks, Lawrence J. and Michael A. Kamins (1988), "The Use of Product Sampling and Advertising: Effects of Sequence of Exposure and Degree of Advertising Claim Exaggeration on Consumers' Belief Strength, Belief Confidence, and Attitudes," *Journal of Marketing Research*, 25 (3), 266-81.

Rutz, Oliver and Randolph Bucklin (2012), "Does Banner Advertising Affect Browsing for Brands? Clickstream Choice Model Says Yes, for Some," *Quantitative Marketing and Economics*, 10 (2), 231-57.

Sherman, Lee and John Deighton (2001), "Banner advertising: Measuring effectiveness and optimizing placement," *Journal of Interactive Marketing*, 15 (2), 60-64.

Sridhar, Shrihari, Murali K. Mantrala, Prasad A. Naik, and Esther Thorson (2011), "Dynamic Marketing Budgeting for Platform Firms: Theory, Evidence, and Application," *Journal of Marketing Research*, 48 (6), 929-43.

Tellis, Gerard J. (1988), "Advertising Exposure, Loyalty, and Brand Purchase: A Two-Stage Model of Choice," *Journal of Marketing Research*, 25 (2), 134-44.

Terlep, Sharon, Suzanne Vranica, and Shayndi Raice (2012), "GM Says Facebook Ads Don't Pay Off," *The Wall Street Journal (*May 15) (accessed June 17, 2012), [available at http://online.wsj.com/article/SB10001424052702304192704577406394017764460.html].

Toupet, A., M. Zhang, K. Rao, S. Varma, and A. Perez (2012), "Humanizing the Internet Cookie? Key Learnings from an Online Panel," in Conference Proceedings from Midwest Association for Public Opinion Research: Chicago, IL.

Wright, Alice A. and John G. Lynch, Jr. (1995), "Communication Effects of Advertising Versus Direct Experience When Both Search and Experience Attributes are Present," *Journal of Consumer Research*, 21 (4), 708-18.

[1] In the web appendix, we present further discussion of the antecedents of this problem based on the data generating process.

[2] Our attempt modeled ad response accounting only for targeting. We found that the signs of the targeting and response parameters were not empirically identifiable due to the data set's low individual level information content, a potential pitfall that Manchanda and co-authors note in their original paper.

[3] When running similar models using sign-up or conversion as the focal behavior, we were unable to obtain significant parameter estimates. This may be due to a lack of effect or to a decrease in observed behaviors caused by moving further through the purchase funnel.

[4] Frequent cookie deletion decreases the probability that both an impression and a site visit would be associated with the same individual. Assuming that those who delete cookies are otherwise similar to those who do not – in line with the findings of Dreze and Zufryden (1998) – including frequently deleted or "short-lived" cookies would incorrectly reduce the correlation between the two events, biasing our parameter estimates towards zero.

[5] One immediate concern may be that logistic regression can underestimate the probability of rare events. Fortunately, the risk of bias declines dramatically as the number of observed events increases, and virtually disappears when the number of observed rare events exceeds 6,000, the upper bound published by King and Zeng (2001). With 16,433 observed site visits, we far exceed this threshold.

[6] We add one prior to log transforming the relevant variables to avoid taking the logarithm of zero.

[7] Note that this expression represents a distribution of probabilities given the data and the joint posterior draws of β; it is not a single probability value.

[8] As an example, consider the case of a one period lag effect (the result we find for our data). Our objective is to compute the lift in probability of a site visit occurring on either the day of exposure or the next day. Thus, we would add the probabilities of site visit for days one and two then subtract the product of those probabilities.

[9] An alternative approach would be to use X in this calculation, with the assumption that future data will be similarly distributed. Given that our data are sparse (i.e., the vast majority of observed covariates are zero), the resulting plots and calculations provide little information beyond simply calculating Equation (3) once using a vector of zeros. We believe that fully enumerating the observed dataset better conveys the range and robustness of the effects.

[10] 98.5% of our observations contain six or fewer impressions.

[11] The observation counts in Table 2 make it unlikely that this null effect is due to a shortage of observations for visitors compared to the other funnel stages.

[12] Because subjects were randomized at the cookie (rather than funnel stage) level, one concern may be that ad effects in later funnel stages are actually cumulative effects from repeated exposures in earlier stages. We have tested this explanation using both model-free and model-based analyses and find no significant differences in estimated effects. Details are available from the authors upon request.

[13] To explore the extent to which our findings on marginal effects and elasticities are driven by functional form, we compare the results of Model 4 to those from a model which replaces the log-transformed impression counts (and associated interaction terms) with linear and squared terms. This specification provides sufficient flexibility to allow differences in marginal effects to arise from the data rather than relying on functional form. We find no significant differences in marginal effects or elasticities.

[14] Details are available from the authors upon request.

[15] The size of the unexposed group will be inversely and exponentially related to the number of funnel stages, which may significantly increase the required control group sample size in analyses with longer purchase funnels.

[16] If this were not done, the approach would recommend that this group be served a negative number of impressions. Also, we again note that the negative elasticity estimate is not significantly different from zero.

[17] This excludes the charity impressions served to the converted customers, as converted customers are not included throughout this comparison.

[18] An alternative approach would be to discard the control group and associated parameters, and then re-estimate the model. This reduces the information content available to estimate the common parameters, while offering no additional benefits. Thus, we believe this approach to be the appropriate comparison.

Table 1
DISTRIBUTION OF OBSERVED IMPRESSION COUNTS

Impression Count	Percent of Cookie Days
0	87.53%
1	5.77%
2	2.39%
3	1.22%
4	0.76%
5	0.49%
6+	1.84%

Notes: This table contains the percentage of observations by observed impression count.

Table 2
OBSERVATIONS BY TREATMENT GROUP AND FUNNEL STAGE

	Firm Ad	Charity Ad	Total
Non-Visitor	4,047,002 (86.62% / 98.30%)	70,082 (92.44% / 1.70%)	4,117,084 (86.71% / 100.00%)
Visitor	281,185 (6.02% / 99.13%)	2,470 (3.26% / 0.87%)	283,655 (5.97% / 100.00%)
Authenticated User	106,690 (2.28% / 98.69%)	1,414 (1.87% / 1.31%)	108,104 (2.28% / 100.00%)
Converted Customer	237,330 (5.08% / 99.23%)	1,847 (2.44% / 0.77%)	239,177 (5.04% / 100.00%)
Total	4,672,207 (100.00% / 98.40%)	75,813 (100.00% / 1.60%)	4,748,020 (100.00% / 100.00%)

Notes: (% Column Total/% Row Total)

Table 3
PARAMETER ESTIMATES & MODEL FIT (BASELINE MODEL)

	Model 1
Intercept	-6.442*** (-6.467,-6.418)
Charity	-0.560*** (-0.814,-0.320)
Visitor	1.874*** (1.831,1.918)
Visitor x Charity	0.522* (0.038,0.985)
Auth. User	2.954*** (2.912,2.996)
Auth. x Charity	0.129 (-0.333,0.565)
Conv. Cust.	2.326*** (2.285,2.365)
Conv x Charity	0.142 (-0.373,0.634)
Log-Likelihood	-98,665 (-98,669,-98,662)
BIC	197,445

Note: This table contains the parameter estimates for the model specification most similar to the model free evidence in Figure 1.

Table 4
PARAMETER ESTIMATES & MODEL FIT

	Model 2	**Model 3**	**Model 4**
Intercept	-6.649*** (-6.676,-6.622)	-6.962*** (-7.012,-6.910)	-6.960*** (-7.012,-6.907)
$\mathbf{ln}(x_{it,f}+1)$	0.851*** (0.825,0.877)	0.820*** (0.795,0.847)	0.869*** (0.842,0.897)
$\mathbf{ln}(x_{it,c}+1)$	0.304` (-0.059,0.594)	0.280 (-0.095,0.586)	0.430* (0.059,0.745)
$\mathbf{ln}(x_{it-1,f}+1)$			-0.159*** (-0.183,-0.134)
$\mathbf{ln}(x_{it-1,c}+1)$			-0.501* (-0.959,-0.113)
Visitor	1.839*** (1.785,1.891)	1.869*** (1.816,1.921)	1.892*** (1.838,1.945)
Auth. User	2.963*** (2.916,3.012)	2.985*** (2.935,3.036)	3.008*** (2.958,3.056)
Conv. Cust.	2.283*** (2.236,2.331)	2.313*** (2.265,2.361)	2.336*** (2.288,2.384)
Visitor x $\mathbf{ln}(x_{it,f}+1)$	-0.408*** (-0.453,-0.362)	-0.403*** (-0.448,-0.358)	-0.389*** (-0.434,-0.345)
Auth. x $\mathbf{ln}(x_{it,f}+1)$	-0.463*** (-0.509,-0.419)	-0.443*** (-0.490,-0.398)	-0.432*** (-0.478,-0.387)
Conv. x $\mathbf{ln}(x_{it,f}+1)$	-0.382*** (-0.422,-0.341)	-0.364*** (-0.404,-0.323)	-0.354*** (-0.395,-0.312)
Visitor x $\mathbf{ln}(x_{it,c}+1)$	0.227 (-0.386,0.768)	0.215 (-0.408,0.785)	0.177 (-0.452,0.746)
Auth. x $\mathbf{ln}(x_{it,c}+1)$	-0.584* (-1.232,0.000)	-0.550` (-1.246,0.059)	-0.524` (-1.210,0.091)
Conv. x $\mathbf{ln}(x_{it,c}+1)$	-1.248* (-2.862,-0.155)	-1.271* (-3.029,-0.128)	-1.268* (-2.935,-0.110)
Week 1		0.590*** (0.534,0.645)	0.610*** (0.556,0.663)
Week 2		0.471*** (0.414,0.527)	0.481*** (0.423,0.539)
Week 3		0.165*** (0.104,0.224)	0.174*** (0.113,0.236)
Week 4		0.034 (-0.027,0.096)	0.042 (-0.019,0.103)
Week 5		0.170*** (0.109,0.230)	0.173*** (0.113,0.234)
Weekend		0.108*** (0.075,0.142)	0.108*** (0.074,0.142)
Log-Likelihood	-96,686 (-96,692,-96,683)	-96,256 (-96,263,-96,251)	-96,169 (-96,176,-96,164)
BIC	193,547	192,774	192,630

*** $\boldsymbol{\alpha} < 0.001$, ** $\boldsymbol{\alpha} < 0.01$, * $\boldsymbol{\alpha} < 0.05$, ` $\boldsymbol{\alpha} < 0.10$

Notes: This table contains the parameter estimates and associated 95% confidence bands based upon our approach.

Table 5
ELASTICITY ESTIMATES BY FUNNEL STAGE

	Median	**90% Confidence Band**	**95% Confidence Band**
Non-Visitor	0.10	(0.06, 0.16)	(0.05, 0.17)
Visitor	0.00	(-0.10, 0.07)	(-0.13, 0.08)
Authenticated User	0.09	(0.05, 0.12)	(0.04, 0.13)
Converted Customer	0.12	(0.07, 0.16)	(0.06, 0.16)

Notes: The second column contains the median elasticity estimate from our model for each funnel stage. The third and fourth columns report the associated confidence bands.

Table 6
ACTUAL IMPRESSION ALLOCATION AND THOSE RECOMMENDED BY THE DORFMAN-STEINER CONDITION

	Actual	**Proposed Estimate**
Non-Visitor	1,364,007	997,960
Visitor	389,757	0
Authenticated User	145,680	869,893

Notes: Column two presents the actual impression allocation by funnel stage, while column three presents the recommended optimal allocation based on the Dorfman-Steiner condition for the proposed approach.

Table 7
EXPECTED NUMBER OF VISITS GIVEN OPTIMAL ALLOCATION RULES

	Existing Impressions		**Incl. Charity Impressions**
	Actual	Proposed Estimate	Actual
Non-Visitor	6,500	6,336	6,511
Visitor	2,914	2,914	2,914
Authenticated User	3,191	4,596	3,196
Total	12,605	13,846	12,621
Quantity Δ		1,241	16
% Change		9.85%	0.12%

Notes: Column three reports the expected number of site visits conditional on the allocations presented in Table 6. The last two columns present a more stringent test, allowing the control group impressions to be served as firm impressions for the actual estimates.

Table 8
COMPARISON OF PROPOSED AND CORRELATIONAL DISTRIBUTIONS

	Effect	**Marginal Effect**	**Elasticity**
Non-Visitor	95.70%	79.40%	87.79%
Visitor	59.85%	24.58%	24.99%
Authenticated User	99.76%	98.74%	99.13%
Converted Customer	99.97%	99.53%	99.59%

Note: This table contains the percent of estimated posterior probability of site visit based on the proposed approach that exceeds the median posterior probability of site visit using correlational approaches.

Table 9
ACTUAL IMPRESSION ALLOCATION AND THOSE RECOMMENDED BY THE DORFMAN-STEINER CONDITION (EXTENDED)

	Actual	**Proposed Estimate**	**Correlational Estimate**
Non-Visitor	1,364,007	997,960	990,463
Visitor	389,757	0	476,979
Authenticated User	145,680	869,893	400,411

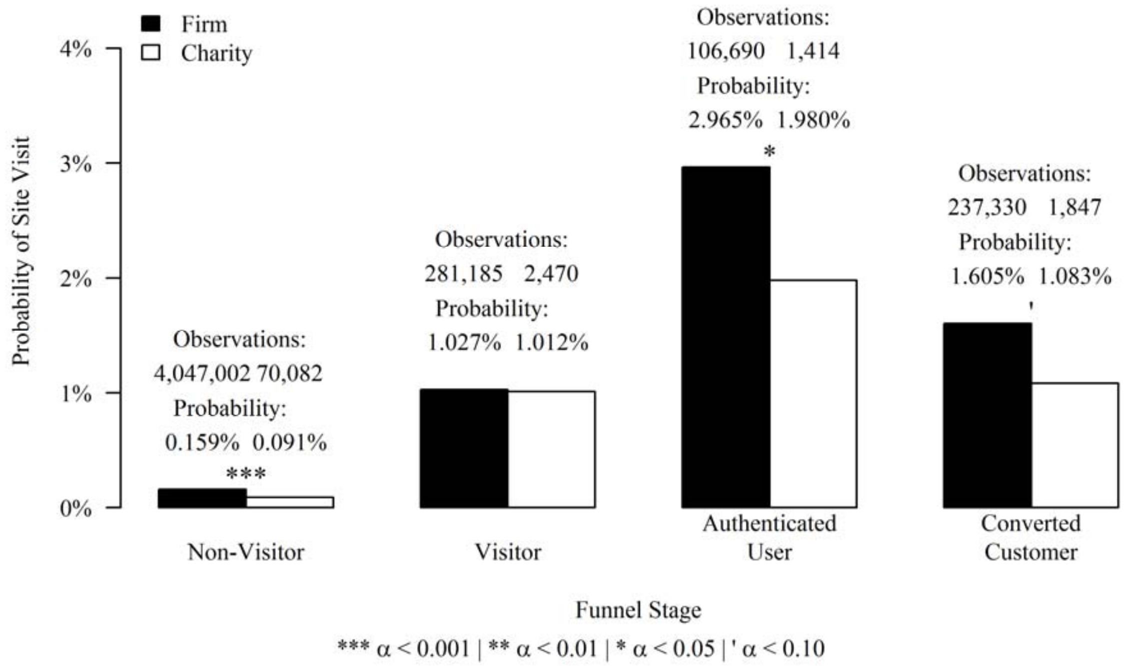
Notes: This table builds on Table 6, adding the recommended impression allocations based on the correlational model estimates.

Table 10
EXPECTED NUMBER OF VISITS GIVEN OPTIMAL ALLOCATION RULES (EXTENDED)

	Existing Impressions Only			Incl. Charity Impressions	
	Actual	Proposed Estimate	Correlational Estimate	Actual	Correlational Estimate
Non-Visitor	6,500	6,336	6,332	6,511	6,341
Visitor	2,914	2,914	2,914	2,914	2,914
Auth. User	3,191	4,596	3,688	3,196	3,701
Total	12,605	13,846	12,935	12,621	12,956
Quantity Δ		1,241	330	16	351
% Change		9.85%	2.62%	0.12%	2.78%

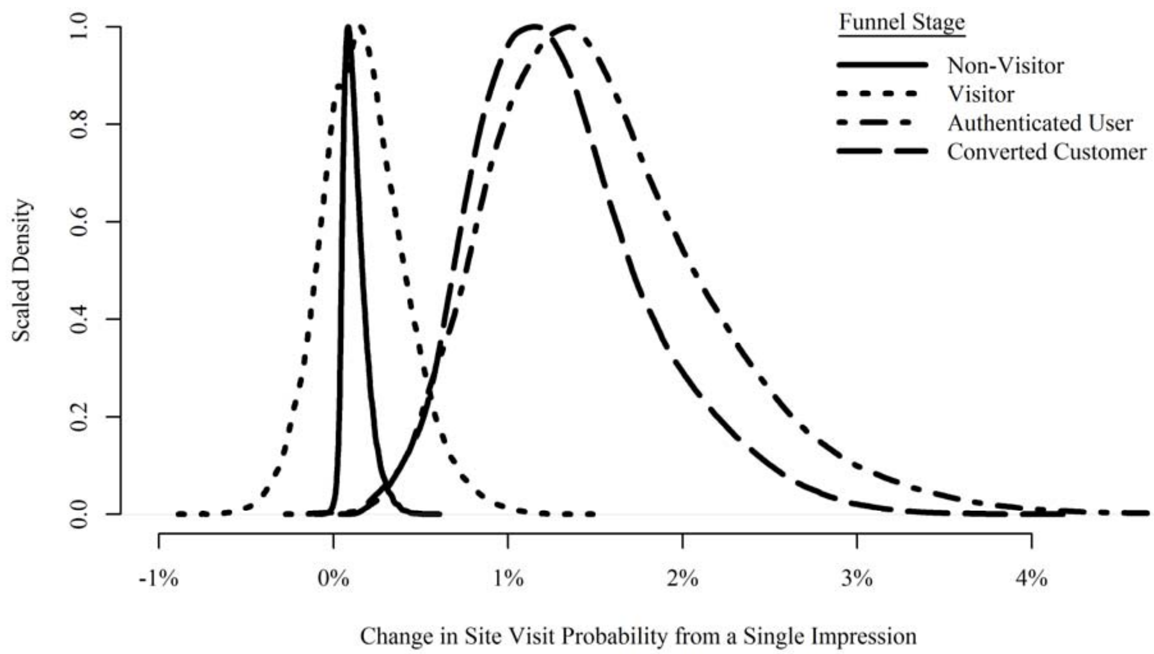
Notes: This table builds on Table 7 by including the expected number of visits given the allocation recommendations resulting from the correlational approach.

Figure 1
PROBABILITY OF SITE VISIT BY FUNNEL STAGE & TREATMENT GROUP



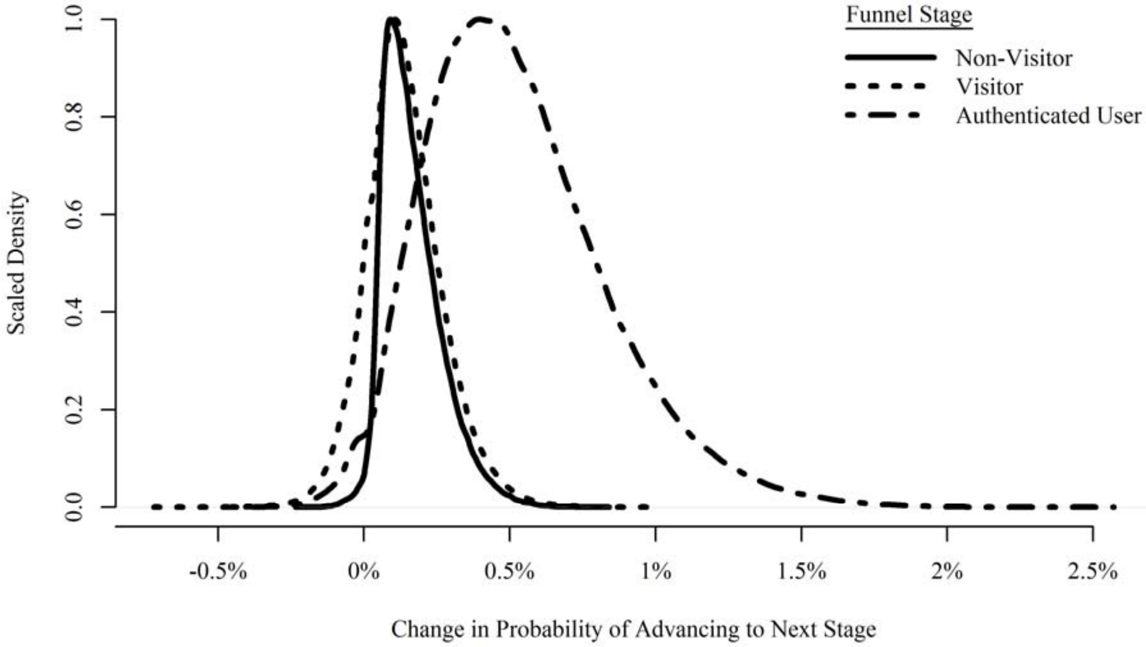
Note: This figure contains the probability of site visit across cookie-day observations, broken down by funnel stage.

Figure 2
EFFECT OF A SINGLE IMPRESSION BY FUNNEL STAGE



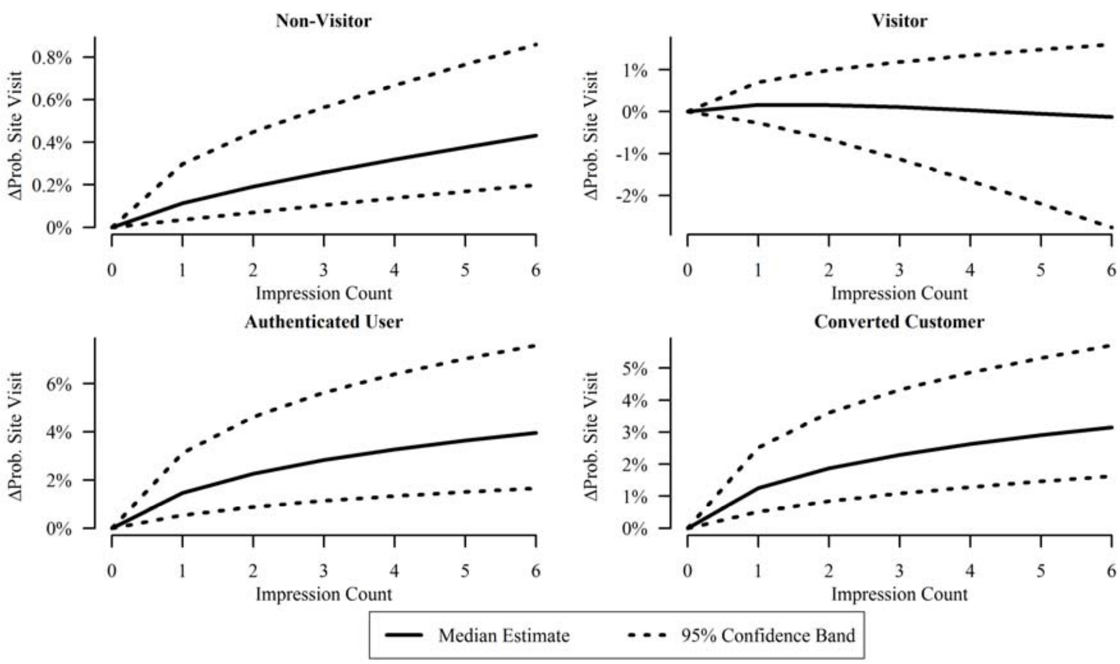
Note: This plot contains the model based effect size estimates by funnel stage.

Figure 3
ESTIMATED EFFECT OF DISPLAY ADVERTISING ON FUNNEL STAGE TRANSITION



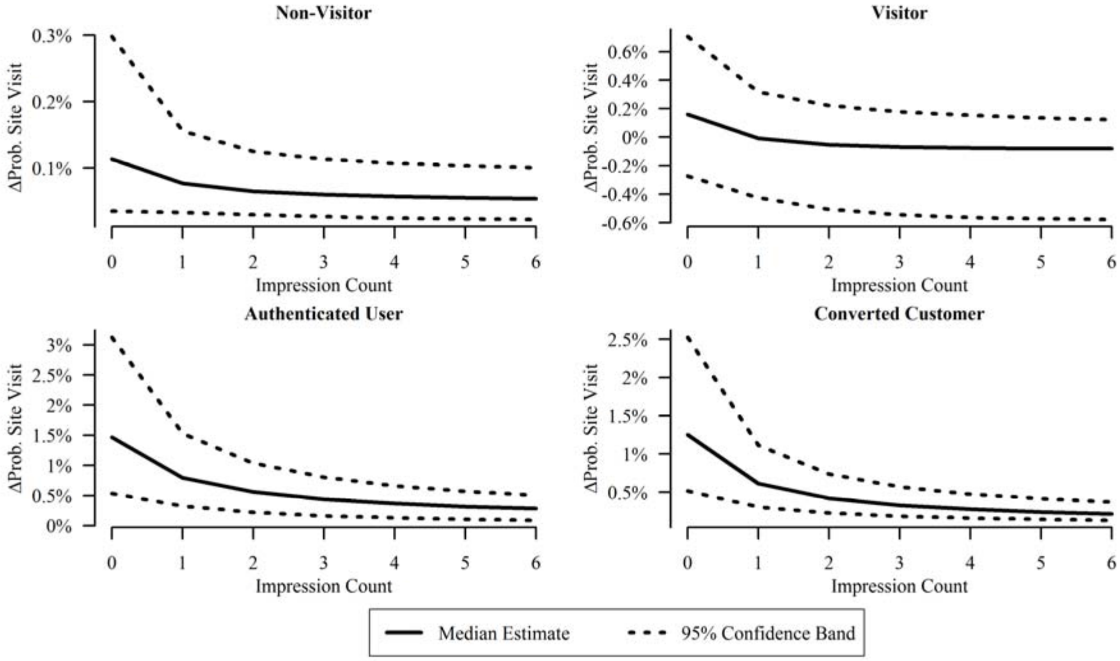
Note: This figure contains the estimated display advertising effect on the transition between funnel stages.

Figure 4
MEDIAN & 95% CONFIDENCE BANDS FOR THE ESTIMATED EFFECTS (1-6 IMPRESSIONS)



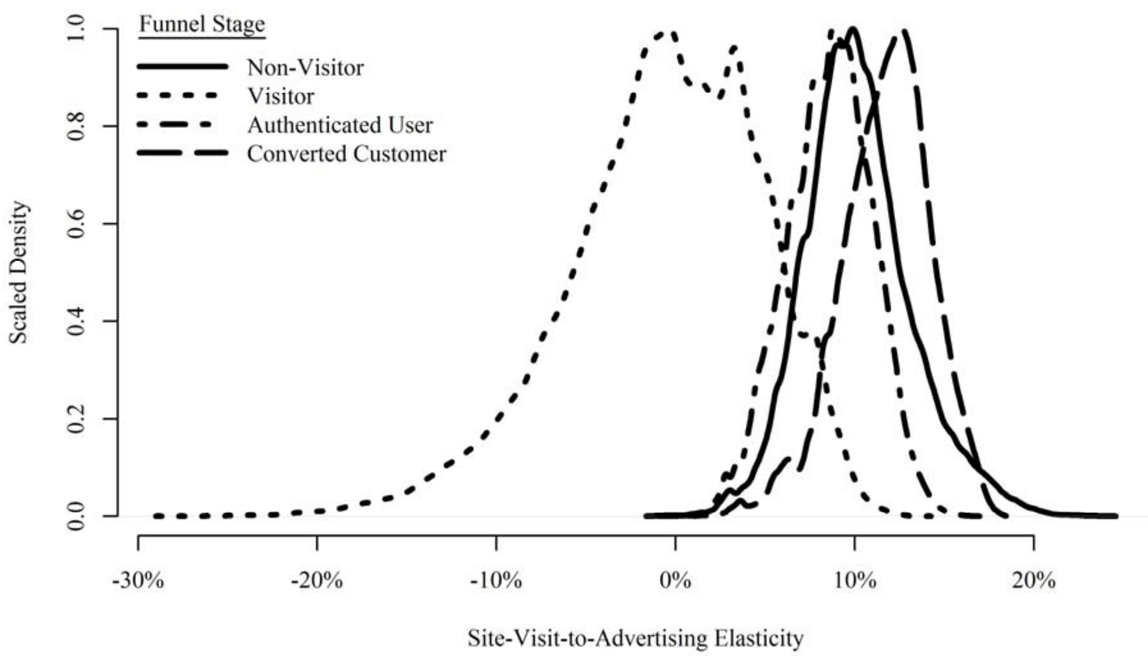
Note: The solid lines reflect the median effect size estimate resulting from our approach, while the dashed lines represent the 95% confidence bands.

Figure 5
MEDIAN & 95% CONFIDENCE BANDS FOR MARGINAL EFFECTS (1-6 IMPRESSIONS)



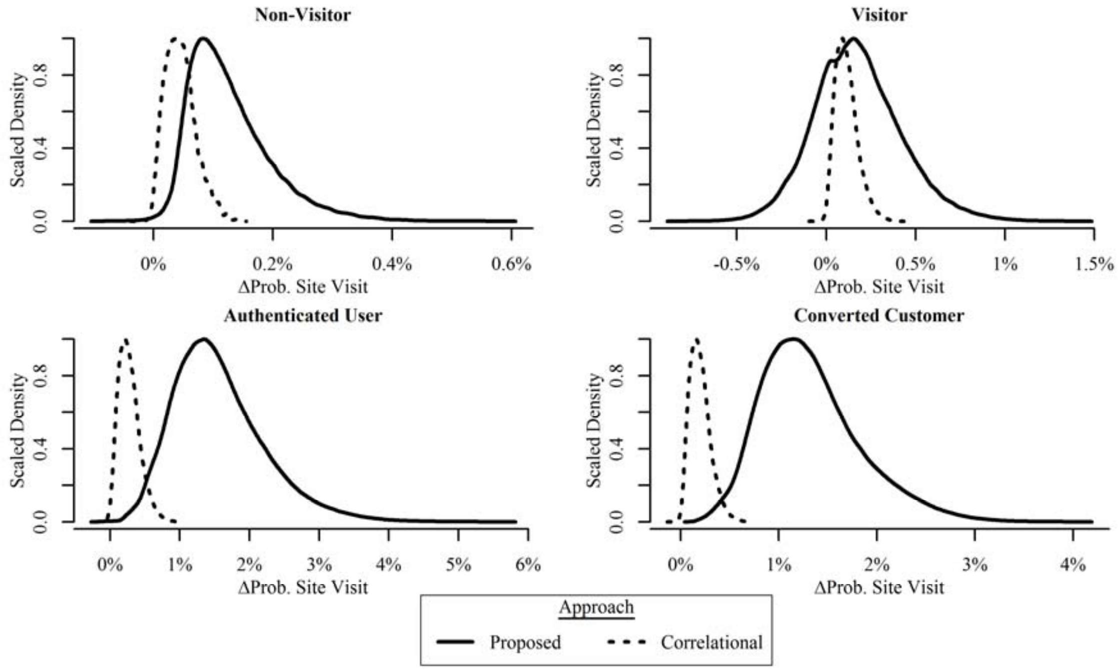
Note: The solid lines reflect the median effect size estimate resulting from our approach, while the dashed lines represent the 95% confidence bands.

Figure 6
ELASTICITY OF MOVING FROM ONE TO TWO IMPRESSIONS



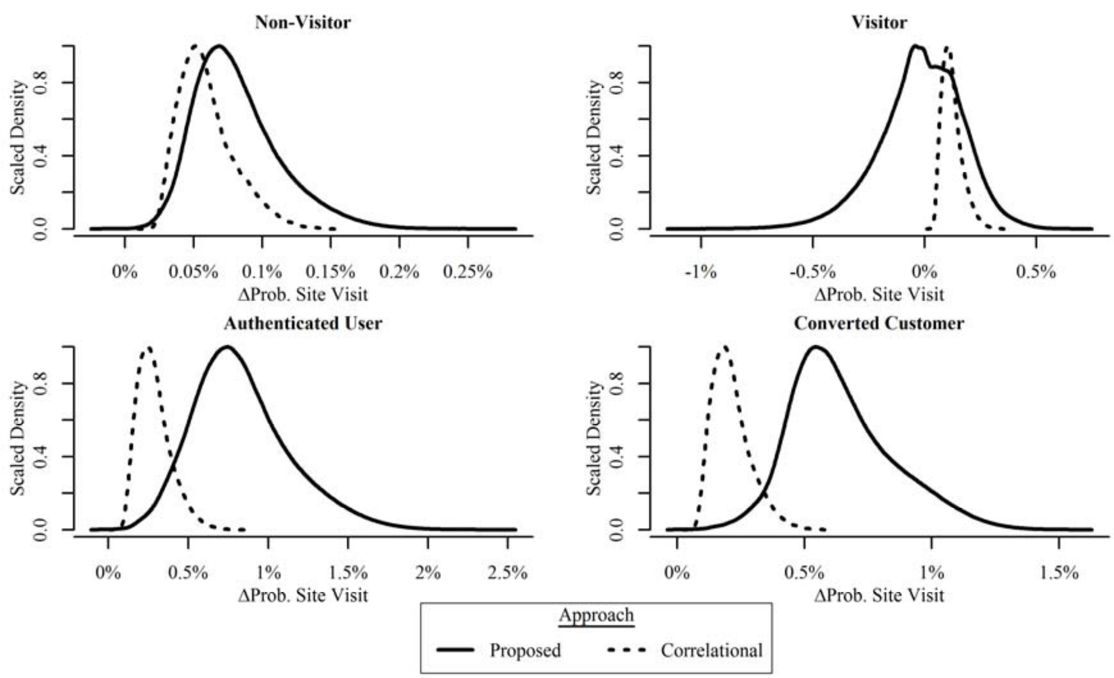
Note: This figure plots the posterior predictive distribution for the elasticity of moving from one to two impressions, broken out by funnel stage.

Figure 7
COMPARISON OF EFFECT OF A SINGLE IMPRESSION



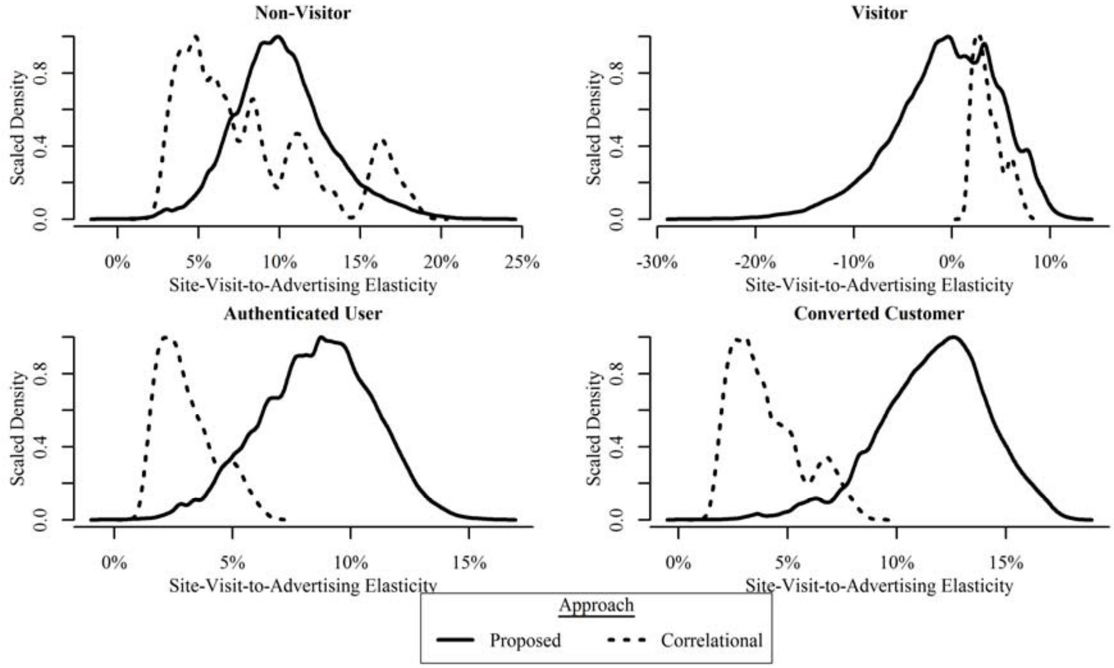
Note: These plots compare the effect size estimates resulting from our approach and a correlational approach for each of the funnel stages.

Figure 8
COMPARISON OF MARGINAL EFFECT FROM ONE TO TWO IMPRESSIONS



Notes: These plots compare the marginal effect estimates resulting from our approach and a correlational appraoch for each of the funnel stages.

Figure 9
COMPARISON OF ELASTICITY FROM ONE TO TWO IMPRESSIONS



Note: These plots compare the elasticity estimates resulting from our approach and a correlational approach for each of the funnel stages.