Does banner advertising affect browsing for brands? clickstream choice model says yes, for some

Oliver J. Rutz & Randolph E. Bucklin

Quantitative Marketing and Economics
QME

ISSN 1570-7156
Volume 10
Number 2

Quant Mark Econ (2012) 10:231-257
DOI 10.1007/s11129-011-9114-3

Springer

Quant Mark Econ (2012) 10:231–257
DOI 10.1007/s11129-011-9114-3

Does banner advertising affect browsing for brands? clickstream choice model says yes, for some

Oliver J. Rutz · Randolph E. Bucklin

Received: 27 May 2009 / Accepted: 12 October 2011 / Published online: 15 November 2011

Abstract This paper investigates how exposure to *Internet display advertising* affects the *subsequent* choices users make of brand-specific pages to view within a website. Using individual-level clickstream data from a third-party automotive website, we tracked the web pages selected by users as they browsed the site and their exposures to premium placement display ads for different vehicle makes (e.g., Ford, Toyota). Pages on the site were classified into those that displayed information about a specific vehicle make (a "make page") versus those that did not (a "non-make page"). For each "make-page" viewed, the specific automotive make selected (e.g., Ford, Toyota) was also recorded. We use these data to develop a model of users' make-specific page choices as a function of prior banner ad exposure on the site. Consumer heterogeneity is captured using a Bayesian Mixture approach. We find that banner ads influence subsequent choices of which make-specific pages to view for ads, served during the current browsing session but not for ads served in previous sessions. The effect of banner ads is also segmented: users in one segment (54%) reacted positively, users in a second segment (46%) were not influenced. Using a standard continuous approach to heterogeneity, we would have concluded–incorrectly–that banner advertising has no effect on the subsequent selection of make-specific pages. For the positively reacting segment, we estimate that the elasticity of make-page choice with respect to banner ad exposure is just under 0.2. Users in this segment appear less focused in their site browsing behavior and tend to stay longer than users in the non-reacting segment.

Keywords Internet · Banner advertising · Clickstream · Logit choice models · Heterogeneity

JEL C01 · C11 · C33 · M31 · M37

O. J. Rutz (✉)
Foster School of Business, University of Washington, Paccar Hall, Seattle, WA 98195, USA
e-mail: orutz@uw.edu

R. E. Bucklin
UCLA Anderson School, 110 Westwood Plaza, Los Angeles, CA 90095, USA
e-mail: rbucklin@anderson.ucla.edu

1 Introduction

Spending on Internet display advertising–or banner ads–in the United States totaled \$6.2 billion in 2010 and made up 24% of the overall spending on Internet advertising of \$26 billion (PricewaterhouseCoopers 2010). This spending is projected to continue to grow much faster than that for offline advertising. A key challenge for banner advertising is that its effectiveness is difficult to evaluate. Two standard metrics for banner advertising, page views and click-through rates, are both problematic. Page views (or impressions) cannot account for whether the consumer actually processes the ad. Industry studies show that, on average, only one out of every 12 banner ads is attended to (Business Week 2006). Click-through rates can discriminate between attended and non-attended ads (e.g., Novak and Hoffman 2000), but the dramatic decline in click-through rates from 7% in 1996 to around 0.2% in 2007 has decreased reliance on the metric (DoubleClick 2003; Business Week 2007). Nevertheless, advertisers continue to increase their spending on Internet display ads and many have announced aggressive plans in this regard (e.g., ComScore 2008).

We seek to contribute to addressing this measurement challenge by leveraging the information in clickstream data. These data allow us to observe the real-time behavior of consumers at websites and also have been previously used to study click-through behavior for banner ads (Chatterjee et al. 2003). Given that click-through rates are quite low (but advertiser interest remains high), we believe that researchers need to develop new ways to aid advertisers in evaluating the effectiveness of their banner ad campaigns. We develop and test one such approach by modeling the detailed, individual-level tracking information contained in the clickstream data collected by web site servers. These data permit analysts to track the page view selections of site visitors, the time spent on each page, and the exposure those users have to display ads served on the site.

In this paper, we investigate the effects of user exposure to premium placement display advertising (also known as exclusive banner ads) at a third-party automotive website. Advertisers do not pay for these ads based on impressions or click-through. Instead, premium display ads are billed similar to traditional print ads: the advertisers buy space on the website and pay a fee for that space (often on a daily basis). These banner ads can be expensive if they are located on prime Internet real estate, e.g., a banner on a leading portal, such as Yahoo! or MSN, can cost up to \$500,000 a day. This is about the same as a 30-second spot on a hit TV series such as CBS's *CSI* (AvenueA/Razorfish 2006). A common goal for exclusive banner ads is brand building as opposed to click-through or sales transactions. Industry analysts report that the premium class of display ads accounts for about two-thirds of the total spending on Internet display ads in the U.S. (ThinkEquity Partners 2007).

The objective of this paper is to use clickstream data–and the browsing behavior it tracks–to determine whether or not banner ads have a short-run effect on user search behavior within the website itself. Specifically, we investigate whether exposure to banner ads alters some of the subsequent page choices a consumer makes at the website. If so, this should give advertisers–and website publishers–confidence that the ads are, at least to some extent, gaining attention, being processed, and influencing consumers to seek out additional information. While these effects are only one short-run aspect of how this advertising might affect shoppers, the ability to induce further information search could be particularly valuable to advertisers in consumer durable categories such as automobiles, appliances, and electronics. For example, sites which could demonstrate these effects for the banner ads placed with them might be able to charge higher rates. Advertisers could use the approach to conduct tests of different creative executions, track wear-out, or aid in media planning and budgeting. The approach could also be more cost effective and less intrusive than other methods such as eye-tracking or real-time surveys.

Apart from click-through rates (e.g., Chatterjee et al. 2003), and to the best of our knowledge, the short-run effects of banner ads have not yet been investigated with clickstream data. In this paper, we do not explore click-through or behavioral measures such as recall or memory. We also do not analyze long-run effects such as changes in brand perceptions or effects on brand choice that could occur long after ad exposure (e.g., Manchanda et al. 2006). Our goal is to show that further clickstream analysis can be one way to address the measurement challenge of online ad effectiveness.

Our focus on the immediate effect of banner ads makes durable goods an especially interesting category. When purchasing durable goods (online as well as offline), such as houses, appliances, or cars, consumers spend considerably more time searching than when buying non-durable goods, such as books (e.g., Klein 1998; Montgomery et al. 2004). Thus, consumers buying a durable good often visit a website repeatedly over multiple days or even weeks before purchasing. As a result, much of the advertising exposure is likely to be well removed from the final purchase occasion. This makes the direct effect of advertising on purchasing difficult to pin down.

In our approach, we track consumers' browsing behavior by using clickstream data, and investigate whether exposure to banner ads changes subsequent page view selections. Following previous research (Moe 2003; Montgomery et al. 2004), we categorized the pages of the website into categories. For the automotive site in our study, any page seen either displays brand-related content (e.g., a Toyota) or other (non-brand-related) content (e.g., financing). This brand/non-brand categorization allows us to examine the sequence of brand-specific page selections made by each individual user. We then model the choice of the current brand-page (e.g., brand A vs. brand B) as a function of previous exposure to banner ads as well as earlier brand-pages.

How best to incorporate unobserved consumer heterogeneity in choice models has been a focus of the marketing literature and an issue that many had considered solved with random coefficients approaches. More recently, the methodological question has been revisited (e.g., Dube et al. 2009, 2010). In particular, a key argument that can be made against random coefficients approaches is the potential for multi-modality. Random coefficients assume that a *single* mean exists around which consumers' preferences are distributed, i.e., the distribution of unobserved heterogeneity is unimodal. Allenby et al. (1998) show using a normal component mixture model that within-component heterogeneity remains substantial and multi-component models represent their data better than a one-component model (i.e., unimodal). In recent work, Dube et al. (2009, 2010) show that unimodal heterogeneity distributions (as implemented in most choice models) fit the data significantly worse than more flexible mixture models. Neither a latent class approach that assumes homogenous consumers within segments nor a random coefficients approach that assumes a uni-modal distribution will properly recover a truly multi-modal continuous heterogeneity distribution.

We investigate this methodological issue as part of our study of banner advertising. Specifically, we use a Bayesian Mixture multinomial logit model and compare the results with a standard Bayesian random coefficients model. A limitation of mixture models is the well-known label-switching problem. We present an alternative approach to this problem in a Bayesian setting that can be implemented directly in the sampling procedure and does not require post-processing or ad-hoc restrictions on the mixture components.

Using the Bayesian mixture approach, we find that the effect of banner advertising is significant for one segment of users, but not a second segment. Had we taken the standard random coefficients approach, the advertising effect would not have been significant at all. In addition to the substantive findings regarding banner advertising, this result suggests that the current effort to develop the mixture approach for use in choice models is quite worthwhile.

The remainder of the paper is organized as follows. We first give background information on online and offline advertising. We then develop our model and describe our dataset. Next, we discuss model selection and the resulting estimates. In conclusion, we discuss the managerial implications of our findings, limitations of our approach, and future research directions.

1.1 Literature

1.1.1 Traditional (offline) advertising

Traditionally, the effect of advertising has been modeled based on aggregate data (e.g., Hanssens et al. 1990; Tellis 2004). The short-run elasticity of advertising is commonly found to be around 0.1 (Tellis 2004). In addition to a short-run effect, advertising can have long-run effects over and above the short-run impact with a half-life of 3 to 4 weeks (Newstead et al. 2009). Thus, much of the return on advertising materializes only in the long-run. For example, one of the most cited advertising studies established that the sales increase during the year of increased advertising is approximately doubled over the next 2 years (Lodish et al. 1995). The effectiveness of some forms of offline advertising and, specifically, TV advertising, has not declined over the last 15 years (at least with respect to sales lift, e.g., Rubinson 2009). This is despite social and technological changes and the development of new advertising channels/instruments. While many studies on offline media have focused on medium- to long-run effects estimation, in our study of banner advertising we will be focusing on short-run effects and specifically investigating how banner ads affect a consumer's search process within a website.

In addition to assessing these effects of advertising, researchers have also sought to understand advertising dynamics in terms of response, copy, wear-in and wear-out, and forgetting. For example, the effect of advertising has mostly been found to be concave (e.g., Hanssens et al. 1990) and the effect of copy, wear-in, and wear-out are important factors (e.g., Naik et al. 1998; Bass et al. 2007). We are also able to investigate some of these factors in our setting. Specifically, we test for concavity in response as well as current versus prior, or lagged, exposure, depending on the user's browsing session in which the banner ads were served.

Due to data limitations, i.e., individual-level advertising data not being available, most of the existing research on the short- and long-run effectiveness of advertising has used methods suited for aggregate data. A notable exception is a recent paper by Terui et al. (2010) who use a choice modeling framework. The authors find that advertising does not directly affect choice, but instead changes the consideration set. This interesting finding enriches the literature on how advertising affects consumer behavior at the micro-level and we seek to at least partially corroborate it in our online setting. Specifically, we focus on how banner ads affect the brand-specific page view choices that consumers make through a website. This allows us to investigate how banner ads affect search for brand information, a process closely related to consideration set formation. Unfortunately, we do not have data on consideration sets nor do we model the effect of banner ads on the user's purchase decision.

1.1.2 Online advertising

The nature of the Internet has opened a wide variety of research streams based on clickstream data, mostly focusing on consumers' browsing and purchasing behavior (e.g., Hubermann et al. 1998; Goldfarb 2002; Moe 2003; Bucklin and Sismeiro 2003; Johnson et al. 2004; Moe and Fader 2004; Montgomery et al. 2004; Sismeiro and Bucklin 2004; Moe 2006a). Turning specifically to the literature on banner advertising, Chatterjee et al. (2003) investigated consumers' direct response (i.e., click-through) to banner ad exposure and found significant heterogeneity in (unobservable) click proneness across consumers. Yet, as noted above, consumers now seldom click on banner ads. Average click-through rates have declined dramatically since the late 1990's, falling to about 0.2% as of 2007 (Business Week 2007). Dreze and Hussherr (2003) investigated why banner ads seem to be ineffective based on click-through rates. They use an eye-tracking device to investigate consumers' attention to online advertising in combination with a survey of Internet users' recall, recognition, and awareness of banner advertising. Their findings suggest that click-through rates are low because consumers actually avoid looking at banner ads–implying that processing of banner ads may be done at the pre-attentive level. Thus, click-through rate might be a poor measure of banner ad performance and traditional measures such as brand awareness and brand recall would be more appropriate.

Manchanda et al. (2006) focus on the relationship between banner advertising and purchase patterns in non-durable goods. They investigate individual purchase timing behavior as a function of advertising exposure and find that banner ads play a role in customer retention via purchase acceleration. In another recent study, Moe (2006b) conducted a field experiment to investigate pop-up promotions. She finds that the characteristics of pop-ups, such as the page on which the pop-up is shown, can be utilized to improve consumer response.

The consumer behavior literature has also focused on exploring the effects of online advertising and has found that many offline effects hold true in an online setting. For example, the duration of ad exposure is an important moderator of banner ad effectiveness (Danaher and Mullarkey 2003). They also find that users in a goal-directed mode are much less likely to recall and recognize banner ads than users who are in a browsing mode. Analogous to offline advertising, Havlena and Graham (2004) find that banner ads have decaying effectiveness for certain categories. For automobiles and pharmaceuticals, a significant decline in brand measures occurs over time. Thus, banner ad effects seem to be short-term in nature in these categories. They hypothesize that these product categories require higher consumer consideration and involvement, making the persuasive power of online advertising of shorter duration and in need of reinforcement over time.

According to Cho and Cheon (2004), consumers avoid looking at advertising on the Internet mainly due to perceived goal impediment. The authors recommend that advertisers should use highly customized context-congruent advertising messages to increase alignment between consumers' goals and advertising. Following up, Moore et al. (2005) investigate the importance of congruity between the website and the ad. They find that congruity has a more favorable effect on attitudes towards the brand, whereas incongruity has a more favorable effect on recall and recognition.

Notwithstanding low click-through rates, the existing literature on banner advertising supports the notion that they can, indeed, be effective. While models have been developed for click-through and for purchase acceleration, there is no clickstream-based modeling approach for assessing the short-run effects of banner advertising on the page view selection decisions made within a web site. Given the insights offered by the various experimental studies from the consumer behavior literature, we believe that further modeling work to study the potential effects of these ads on various pre-purchase decisions (such as what to search for next) is a worthwhile research endeavor.

2 Modeling approach

2.1 Modeling page view choice data for durable goods

The web pages of most any online retailer selling durable goods can be categorized into two basic categories: brand-pages, which display a brand, and non-brand-pages, which display other, non-brand-related information. A brand page, for example, could display a plasma TV from Sony, whereas a non-brand-page could display information on shipping. In our approach, we model the consumer's brand-page choices (i.e., which brands are searched), with the goal of investigating the effect that banner ad exposure has on the consumer's search process.

Given our focus on user page view choices, we note that our work is related to the browsing path study conducted by Montgomery et al. (2004). Their study focused on understanding purchase conversion based on browsing behavior for books and CDs. In their model, a consumer who exhibited focused browsing, i.e., limited page-views and a direct route to a certain product, was found more likely to buy. For non-durables, the notion of a browsing vs. purchase goal is intuitive and predictive. However, for a durable goods purchase such as a new car, information search is considerably longer. In our data, for example, consumers who ultimately buy spent significantly more time on the site, viewed more pages and returned more often than consumers who did not. Thus, it is not clear which types of browsing behavior might be most predictive of purchase when those decisions occur after multiple sessions, each with a significant number of page views. Because we study a durable goods category (new cars) and focus on the short-term pre-purchase effects of banner ads, we do not follow the Montgomery et al. (2004) approach in this study. Instead, we investigate brand-page choices as a function of previous brand and banner ad exposures from the current and preceding sessions.[1] When the consumer comes to a page that requires a brand-page choice to proceed, we model the probabilities that the consumer selects the brand-page corresponding to each alternative (e.g., does the consumer request a page that shows a Toyota, a BMW or another car brand). In this fashion, we seek to capture the potential effects of prior banner ad exposure,[2] but do so in a parsimonious framework which can be implemented on a widespread basis.

2.2 Covariates

We investigate whether current and previous within-site banner ad exposure is predictive of the brands a consumer searches. Recall that Dreze and Hussherr (2003) found that consumers try to avoid looking at banner ads. Nonetheless, consumers in their study still looked at approximately 50% of the banner ads they were served. If a primary reason for consumers to avoid looking at banner ads is perceived goal impediment (Cho and Cheon 2004), content congruity with the banner ads could turn out to be a critical factor (Moore et al. 2005). Our website served only banner ads for car brands, providing a high level of congruity between the ads and the site. Hence, perceived goal impediment should not be an important antecedent to banner ad avoidance for most users in our setting.

Next to ad exposure by itself, the recency of the ad exposure could also matter. For example, ad exposure from a previous web site session might be less likely to influence current page view choice behavior. Existing research shows that banner ads exhibit decaying effectiveness for automobiles (Havlena and Graham 2004). In their study, a significant decline in brand measures occurs over time, and the effects of advertising seem to be mainly short-run in nature. We investigate whether this effect holds in our data, i.e., does exposure to banner ads in previous sessions alter the brands a consumer searches in the current session? We also examine whether the

[1] Potentially, these choices might not be independent. For example, consumers could be more likely to view finance pages if they have previously viewed a research page. We created page view profiles of users to test whether certain page types (e.g., research, finance, car buying guides) tend to appear together. We found no evidence for this; further details are available upon request.

[2] Consumers exposed to many ads could have different page view profiles on this site. This could indicate that ad exposure not only changes the brands viewed, but also the path taken through the site, which we are not modeling. We compared the page view profiles (i.e., the mix of different page types) of consumers with low banner ad exposure to consumers with high ad exposure and found no differences. Thus, we conclude that ad exposure does not affect the actual browsing path as it relates to page-type choice. Further details are available on request.

existing finding that duration is a better measure for banner ad exposure than count (e.g., Danaher and Mullarkey 2003) can be confirmed in our data.

Our clickstream data provide information on a user's prior exposures to both web pages as well as banner ads. We use the information on previous brand-page exposure to control for consumer interest and browsing state dependence. Previous brand-page exposure should reveal unobserved consumer preferences in the search process. For example, a consumer who has primarily requested pages pertaining to Chevrolet may be more interested in this brand than other brands. Including this variable in the model should capture differences in page view choice probabilities both across individuals and over time. After controlling for brand-page exposure as a covariate, we will use the Bayesian Mixture approach to capture remaining unobserved heterogeneity.

2.3 Model specification

Existing research on browsing behavior (e.g., Hubermann et al. 1998; Sismeiro and Bucklin 2004) has conceptualized browsing as a repeated discrete choice process. In general, a consumer has the option to stay on the current page, to navigate to another (linked) page or to end his session. In our approach, we model the brand-page choices made by each user. Note that we do not model the choices of other page types or whether the consumer decides to end his session. This is because we are focused on understanding the effects banner ads have on the user's desire to search for more brand specific information. From a behavioral perspective, we would not expect that (branded) banner ad exposures influence the choice of web pages that are unrelated to brand-specific information[3] or change the probability of ending the session.[4] In other words, it seems unlikely that a banner ad for a Toyota would make a consumer more likely to navigate to a financing page versus a banner ad for GMC. Thus, conditional on the consumer's decision to view a page containing brand-specific information, we model which brand the consumer selects. We are interested in understanding the effect banner advertising has on a visitor's propensity to search an advertised brand versus other brands. In using the conditional modeling approach, we do not investigate whether advertising exposure makes a consumer more likely to continue searching.

To estimate our model, we take each individual's browsing data and convert it into a timed sequence of brand-page choices. We consider only those browsing choices that result in the consumer visiting a page that shows a brand. Browsing choices that do not result in exposure to a brand are excluded from our analysis. For example, assume individual A starts his session on the company homepage. From this page he can navigate to pages that show either brand A or brand B

[3] We examined whether banner advertising affects page-type choice using page view profiles. We found no evidence for this in our data.

[4] Banner advertising exposure could shorten or prolong the browsing path. We have tested whether this occurs in our data by estimating a version of the binary site exit model proposed by Bucklin and Sismeiro (2003). We find no evidence that banner ad exposure affects the user's decision to exit the site. This holds when ad exposure is measured either by counts of banner ads or by duration of banner ad exposure. Further details are available on request.

(first brand-page choice, $t=1$). Let us assume he decided to navigate to the page that shows brand A, i.e., he chose a brand-page showing brand A. Next, he navigates to pages that do not show brands, e.g., pages that show shipping or payment information; these page choices do not enter the model. At one of these pages, however, he is exposed to a banner ad for brand A. He then navigates back to the homepage and from the homepage he chooses to look at brand B (second brand-page choice, $t=2$). Subsequently, he ends the session. This visitor made two brand-page choices ($t=1$ and $t=2$). At $t=1$ he had no previous exposure to brands and no exposure to ads. At $t=2$ he had a previous exposure to brand A as well as a previous exposure to an ad for brand A. We model his brand-page choices at $t=1$ and $t=2$ based on previous brand as well as previous ad exposures. For users with multiple sessions, we also use the brand and ad exposure data from prior sessions in our model of current brand-page choice.

Formally, we specify a multinomial logit choice model[5] for a consumer's brand-page sequence via the following choice utility $u_{i,t,j}$

$$u_{i,t,j} = c_{i,j} + \sum_{k=1}^{K^{Brand}} \alpha_{i,k} f\left(Brand_{i,t,j,k}\right) + \sum_{k=1}^{K^{ad}} \beta_{i,k} f\left(Ad_{i,t,j,k}\right) + \varepsilon_{i,t,j} \tag{1}$$

where i is the individual, t is time (i.e., count in the browsing sequence), j is the brand, k is the session ID ($k=1$ is the current session), $c_{i,j}$ is a individual specific brand intercept, $\alpha_{i,k}$ and $\beta_{i,k}$ are the individual-specific response coefficients to brand and ad exposures from session k, $Brand_{i,t,j,k}$ is the number of brand j exposures from session k for individual i at time t, $Ad_{i,t,j,k}$ is the exposure measure to ads for brand j from session k for individual i at time t, f describes the functional form with which the exposures measures enter (linear, quadratic or logarithmic), and $c_{i,j}$, $\alpha_{i,k}$, $\beta_{i,k}$, K^{Brand}, and K^{ad} are parameters to be estimated.

We estimate alternative model specifications along two dimensions. First, we test for *functional form* (f) of past exposure (linear vs. quadratic vs. logarithmic) to brands and banner ads. Second, we test for *measurement* of banner ad exposure in terms of ad counts versus exposure duration.

2.4 Consumer heterogeneity and estimation

We use a Bayesian Mixture approach to capture heterogeneity and estimate the model by using Monte Carlo Markov Chain (MCMC). This approach combines the best of two worlds in that we obtain the classification power of a latent class approach and the individual-level estimates of the standard continuous Bayesian approach. Previous research (Andrews et al. 2002) has shown that, depending on the data, it is not clear whether to prefer a discrete or a continuous approach to capturing heterogeneity. A latent class approach might underestimate diversity in preferences

[5] Given that we model a browsing path, autocorrelation could be a potential concern. We have tested this by using a multiperiod multinomial probit (MMP, see McCulloch and Rossi 1996 or Geweke et al. 1997 for details) that allows for an AR(1) process in the error term. We find no evidence for autocorrelation in our application. For further details, please contact the authors.

across consumers by assigning them to segments and incorrectly defining all consumers within a segment as homogeneous. A continuous approach is incorrect if the underlying population distribution is not unimodal. For example, it is common to assume a Gaussian prior distribution to model the population distribution of the individual-level parameters β_i, i.e., $\beta_i \sim N(\mu_b, \Sigma_b)$. If β_i stems from a bimodal distribution–i.e., some consumers are influenced by advertising and some are not–the assumption of a unimodal prior will bias the results. This bias comes from the fact that a unimodal prior is informative–it follows that the posterior distribution is unimodal as well. A better approach allows for a multimodal prior and, in turn, a multimodal posterior in the absence of information pointing towards a unimodal distribution.

From the utility specification in Eq. (1) and the assumption of an extreme value error, the choice probability of individual i searching for brand j at time t is given as

$$P_{i,j,t} = \left(\frac{\exp\left(u_{i,j,t}\right)}{\sum_k \exp\left(u_{i,k,t}\right)} \right), \tag{2}$$

and the likelihood of the model is given as

$$Likelihood = \prod_i \prod_j \prod_t \left(P_{i,j,t}\right)^{I(j=k,t)}. \tag{3}$$

In a "traditional" (i.e., in the marketing literature) Bayesian treatment of heterogeneity, one assumes that the set of individual parameters $\theta_t = [c_i \ \alpha_i \ \beta_i]$ (for simplicity of notation j and k subscripts are omitted) come from a common (typically) unimodal population distribution, such as the Gaussian: $\theta_i \sim N(\mu_\theta, \Sigma_\theta)$ with hyper-priors $\mu_\theta \sim N(\mu_0, \Sigma_0)$ and $\Sigma_\theta \sim Wishart(\nu_1, \nu_2)$. In a Bayesian Mixture Model the assumption of a unimodal population distribution is relaxed, allowing for multimodal population distributions. This can be implemented by using a mixture of distributions as the population distribution (e.g., Diebolt and Robert 1994; Allenby et al. 1998). In the case of a choice model (and keeping with the common Gaussian set-up) we model the individual parameters θ_i as a mixture of Gaussians:

$$\theta_i \sim \sum_{s=1}^{S} \pi_s N(\mu_s, \Sigma_s), \tag{4}$$

where $\pi = (\pi_1, \ldots, \pi_s)$ are the *mixture proportions* (which are constrained to be non-negative and to sum to 1) and $N(\mu_s, \Sigma_s)$ are the different mixture components, where $\mu = (\mu_1, \ldots, \mu_S)$ and $\Sigma = (\Sigma_1, \ldots, \Sigma_S)$. We use the standard (uninformative) Gaussian hyper-priors $\mu_s \sim N(\mu_0, \Sigma_0)$ and $\Sigma_s \sim Wishart(\nu_1, \nu_2)$. We can conceptualize this approach as assuming that each observation θ_i (i.e., the individual parameters) arises from an unknown component z_i of the mixture, where $z_i, \ldots z_n$ are realizations of the independent and identically distributed random variables $Z_i, \ldots Z_n$ with a probability mass function

$$\Pr(Z_i = s | \pi, \mu, \Sigma) = \pi_s \ , \ \ (i = 1, \ldots, n \text{ and } s = 1, \ldots, S). \tag{3}$$

Or, in other words, Z_i are i.i.d. draws from the multinomial distribution:

$$Z_i|\pi, \mu, \Sigma \sim multinomial(1, \pi). \tag{4}$$

Conditional on the Z_s, θ_i is an independent observation from the density

$$p(\theta_i|Z_i = s, \pi, \mu, \Sigma) = N(\theta_i|\mu_s, \Sigma_s.). \tag{5}$$

A Bayesian approach requires the specification of a prior distribution of p (π, μ, Σ) for the parameters of the mixture model. Inference is based on the posterior distribution $p(\pi, \mu, \Sigma|\theta)$ and the quantities of interest are calculated by integrating out the model parameters. The Dirichlet distribution is a flexible and computationally convenient alternative for modeling parameters such as π that lie between 0 and 1 and sum to 1:

$$\pi \sim Dir(\rho), \tag{6}$$

where ρ is a S-vector of prior hyperparameters $\rho = (\rho_1,\ldots,\rho_S)$. Please see the Appendix for a detailed description of the sampler.

2.5 Label switching

A common problem with mixture models is label switching. For any permutation ξ of [1…S], the corresponding permutation of the vector (π, μ, Σ) is given by

$$\xi(\pi, \mu, V) = \left(\left(\pi_{\xi(1)}, \ldots, \pi_{\xi(S)}\right), \left(\left(\theta_{\xi(1)}, V_{\xi(1)}\right), \ldots, \left(\theta_{\xi(S)}, V_{\xi(S)}\right)\right)\right). \tag{7}$$

Label switching is a problem as the mixture likelihood

$$L(\mu, V|\theta) = \prod_{i=1}^{n} \left[\pi_1 N(\theta_i|\mu_1, V)_1 + \ldots + \pi_S N(\theta_i|\mu_S, V_S)\right] \tag{8}$$

is invariant for all permutations of (π, μ, Σ). Thus, when using the draws to recover the posterior means of the component-specific parameters (i.e., μ and Σ) we get nonsensical results, as label switching does not allow us to correctly identify the specific components.

The statistics literature has proposed a variety of methods to address the label switching problem inherent in mixture models (e.g., Celeux et al. 1996; Richardson and Green 1997). A commonly used method, imposing an ordinal restriction on the components masses (e.g., Allenby et al. 1998), will be ineffective if it does not remove the symmetry in the posterior distribution. We use the Relabeling algorithm proposed by Stephens (2000) which provides a simple and robust method to address label switching. It is based on a decision-theoretic approach that applies a loss function that is invariant under permutation of the parameter vector. This property allows relabeling of the MCMC draws to correct for label switching (for a detailed discussion see Stephens 2000).

Let $Q = [q_{is}]$ be an $(n \times S)$ matrix that contains the probability that observation i (i.e., μ_i, Σ_i) is assigned to component s. Let $P(\mu, \Sigma)$ be the matrix of classification probabilities $[p_{is}(\mu, \Sigma)]$, where:

$$p_{is}(\mu, V) = \Pr(Z_i = s|\theta, \pi, \mu, V) = \frac{\pi_s N(\theta_i|\mu_s, V_s)}{\sum_k \pi_k N(\theta_i|\mu_k, V_k)}. \tag{9}$$

Initialize $\widehat{Q}^0$by using a small preliminary MCMC sample (Celeux 1998) in which it is assumed that no label switching occurs. At stage t in the Sampler we have a current draw of the mixture parameters (μ^t, Σ^t) and we have a current estimate of $\widehat{Q}^{t-1} = \left[\widehat{q}_{is}^{t-1}\right]$. Choose a permutation $\xi_t^*$to minimize

$$\xi_t^+ = \arg\min_{\xi_t} \sum_{i=1}^{n} \sum_{s=1}^{S} p_{is}(\xi_t(\mu^t, V^t)) \log\left[\frac{p_{is}(\xi_t(\mu^t, V^t))}{\widehat{q}_{is}^{t-1}}\right]. \tag{10}$$

Relabel the draws in accord with permutation ξ_t^*. Calculate the next $\widehat{Q}^t$as

$$\widehat{Q}^t = \frac{t\widehat{Q}^{t-1} + P\left(\xi_t^*(\mu^t, V^t)\right)}{t+1}. \tag{11}$$

This relabeling algorithm can be implemented in the proposed sampler by adding two steps (see Appendix).

3 Data

Our data come from a major commercial website in the automotive industry; the company wishes to remain anonymous. Consumers can view information about the company and the site, research vehicle information, and compare and configure almost all commercially available cars and light trucks. An order for a new car can be initiated online (known as submitting a lead). Our dataset contains consumers who have submitted leads. The pages of the site can be categorized into two basic categories: make[6]-pages, which display a make, and non-make-pages, which display other, non-make-related information such as financing or buying guides.

To view information about a given make on the website, the consumer must request it; we call this action a make-page choice. The site is structured so that only certain pages allow the consumer to choose a make-page. The consumer must always return to one of these pages before he or she can execute a make-page choice. Thus, we are able to pinpoint the location/time of each make-page choice in each user's browsing path. Our proposed model enables us to investigate how these make-page choices are influenced by previous exposure to banner advertising on the site.

[6] In the automotive industry a brand is generally referred to as a make, i.e., Toyota is a make. For the reminder of the paper we will adhere to that definition and use "make" instead of "brand"

3.1 Data collection

The company collected the data using a proprietary, clickstream tracking system. For each page, the following information was recorded: page ID, cookie ID, day and time of request, and the type of page, e.g., displaying a make vs. financing page. In addition to the cookie-level browsing information, we have information on the site's premium display ads, or so-called exclusive banner ads.[7] These ads are displayed exclusively on a specific page, e.g., the *New Car Search* page, for a certain amount of time, e.g., 2 weeks. Exclusive ads can be displayed on make- and non-make-pages alike. Placement of these ads was independent from the make displayed on make-pages. For example, a page showing information about a Ford can have an ad for a Toyota on it. Because the ad serving was "hard wired" in this fashion, ad exposures should be independent from the make-page information the consumer has requested. We are able to pinpoint which ads a given consumer has been exposed to, when the exposures occurred, and for how long the exposures lasted. We measure the duration of ad exposure by the time until the user's next page request.

In addition to the premium ads, the website also displays other banner ads, which are randomly displayed on non-exclusive pages and have not been recorded by the system. These are sometimes referred to as "remnant" display ads. While it would no doubt be of interest to model the effect of these ads alongside the premium ones, we believe that their absence will not lead to biased results. Because the non-exclusive ads are displayed randomly, they should, in expectation, affect consumers to an equal extent. Consumers in our sample also average about 100 page views. The high number of page views should help to ensure that any small numbers effects wash out across consumers. Accordingly, we focus on the incremental effect that the premium or exclusive ads have on the page choice decisions of consumers.

3.2 Browsing path example

To clarify the structure of the data, we discuss an illustrative browsing path (see Fig. 1). Our illustrative consumer starts his session on the *Home* page, where he executes his first make-page choice by requesting make 1 (the *Home* page is a non-make page). The consumer is exposed to an ad for make 3 for 12 s on the *Home* page. The next page is the *ZIP code* page, where he is required to enter his ZIP code (the *ZIP code* page is a non-make page). The consumer again is exposed to an ad for make 3, this time for 20 s. So far, the consumer has been exposed to ads for make 3 for 32 s in the current session. Next, he encounters his first make-page, the *Car Configuration* page. At this point, he has now seen his first make of the session, make 1. There is no ad exposure on this page. He continues to the *Finance* page for the next two clicks. The *Finance* page is a non-make page and there is no ad exposure. He then ends his session. To recap, in

[7] In our case, the advertiser is buying a certain amount of exposure and the website is serving these ads so that the exposure criteria are fulfilled. This is a practice still in widespread use.

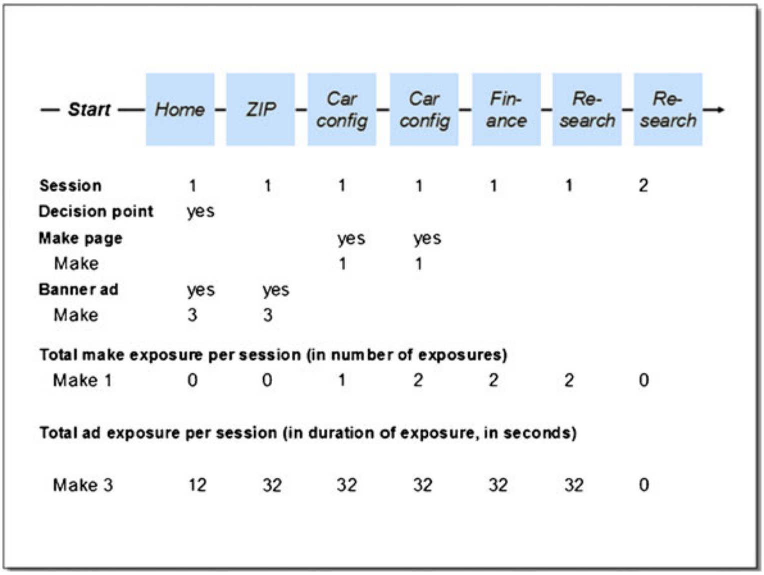

Fig. 1 Illustrative browsing path

session 1, the consumer made one make-page choice and, as a result, saw one make (make 1) and was exposed to two ads (for make 3) with a total duration of 32 s. This information will be used in his second session as the previous session data where we model whether the make and ad exposure in session 1 has an effect on the make-page choice in session 2.

3.3 Sample details

Our data are from February 2004 (29 days). We restricted the sample to all consumers who submitted leads during the last 2 weeks of February (14th to 29th). This gives 4,689 consumers who had 376,960 page views. This approach enables us to capture most, if not all, of the browsing that was done prior to a lead submission. To assess this, we compared the total number of page views in February for consumers who submitted leads during the *first* week of our sample (14th to 21st) with the total number of page views in February for consumers who submitted leads during the *second* week of our sample (22nd to 29th, see Table 1). In addition, we investigated page views for these groups of consumers for one particular day in both weeks and the split of page views over the different car categories on that day (see Table 1). We found no substantial differences in the number of page views between either week or in any of the other comparisons. The results of this analysis indicate that most of the consumer's browsing is done in the 2 weeks before the consumer submits a lead.

The sample contains nearly all automobile makes and models available (see Table 2). We follow the established practice in choice modeling and focus on the major makes in the sample. We retained all makes with a sample market share

Table 1 Leads and page views by period and vehicle category

Period	From	To	# Days	# Leads	# Page views
1	2/14/2004	2/21/2004	8	2,375	199,566
2	2/22/2004	2/29/2004	8	2,314	177,394
3	2/14/2004*	2/14/2004	1	290	20,312
4	2/28/2004*	2/28/2004	1	284	21,078
	Number of Leads				
Period	Car	SUV	Truck	Van	Total
1	1,007 (42.4%)	857 (36.1%)	314 (13.2%)	197 (8.3%)	2,375
2	996 (43.0%)	819 (35.4%)	318 (13.7%)	181 (7.8%)	2,314
3	112 (38.6%)	109 (37.6%)	42 (14.5%)	27 (9.3%)	290
4	112 (39.4%)	91 (32.0%)	46 (16.2%)	35 (12.3%)	284

* Both days are Saturdays

greater than 1.5%, leaving a total of 10 major makes (see Table 3). Based on these data, we then took a random sample of 250 individuals. This resulted in 26,087 page views, of which 7,690 displayed one of the 10 major makes (29.5%). In order to view the 7,690 make pages, the 250 individuals made 4,280 make choices.[8] The average number (per consumer) of page views was 104.3, and the average number of major make-page views 30.8 (see Table 4). The consumers were exposed to 6,073 ads (1,608, or 26.5%, for Ford (make 3), 2,110, or 34.7%, for GMC (make 4), and 2,355, or 38.8%, for Toyota (make 10), see Table 4). This leads to an average exposure per consumer of 24.3 ads (6.4 for Ford, 8.4 for GMC, and 9.4 for Toyota, see Table 4).

For each consumer, we created the individual sequence of page views including session information and page duration information. Following established practice in clickstream analysis, we define a session as a sequence of page requests that have no more than a 10-minute break. The page duration is measured in seconds. Our data therefore allow us to test for the effect of banner ad exposure in terms of both number of ads as well as duration.

4 Empirical results

Using the general model specified above, we test whether advertising effects differ across consumers by modeling heterogeneity through a Bayesian Mixture approach. In accord with established findings, we test whether banner advertising exposure exhibits a diminishing marginal effect by varying the functional form by which it enters the model (specifically, we test log, linear, and quadratic formulations). We also test for the decay effect of advertising by varying the

[8] Note that a consumer can look at multiple pages with information on the same car after a make choice. For example, pages can display pictures of the car from different angles or provide information on extras available for the car.

Table 2 Leads, makes, and models by vehicle category

Category	Segment	# Leads	# Makes	# Models
Car	Compact	683	16	26
	Midsize	800	16	31
	Luxury	392	14	58
	Sporty	116	10	13
SUV	Compact	454	12	14
	Midsize	672	16	21
	Fullsize	354	5	16
	Luxury	179	11	19
Truck	Compact	209	7	9
	Fullsize	467	6	17
Van	Mini	336	12	14
	Fullsize	27	3	5
Total		4,689	43*	243

* Brands compete in different categories and segments

number of sessions (K^{ad}) from which we include past ad exposures. (We also test for the decay effects for state dependence by varying past make exposures, i.e., K^{make}.) Lastly, we examine whether user exposure to ads is best measured by count or duration.

4.1 Summary of key results

Our initial empirical results yielded a series of interesting findings. First, the two-component Mixture provides the best fit (see Table 5). A traditional Bayesian approach to heterogeneity using a unimodal Normal prior fits the data significantly worse.[9] The traditional approach would have led us to the erroneous conclusion that banner advertising exposure has no effect on make-page choice because the advertising coefficient is not significant in the case of a unimodal prior.[10] Allowing for a multimodal prior, we find that there is a segment of consumers that is affected by banner advertising when it comes to subsequent make-page choice decisions (see Table 6). Second, we find diminishing returns (i.e., the logarithmic functional form provides the best fit) to the amount of ad exposure, e.g., the 2nd exposure will have a bigger influence on make-page choice than the 12th exposure. This finding is consistent with established results in advertising research—the response to advertising is mostly concave (e.g., Tellis 2004). Third, current and previous session make exposures are predictive of current make-page choice, while exposures dating further back do not have a significant effect. Thus, we conclude that $K^{make} = 2$. Fourth, only current banner ad exposures are predictive of current make-page choice, while advertising exposures from the previous sessions have no significant effect. Accordingly, we conclude that $K^{ad} = 1$. Fifth, the effect of banner ads is better

[9] Note that Allenby et al. 1998 find a similar drop in Log Marginal Density (LMD) when comparing a latent class model to a Mixture-Component Model.

[10] Mean of 0.012 with a 95% coverage interval (−0.097, 0.104).

Table 3 Leads for makes included in the estimation sample

Make	Make ID	# Leads	Percent
Chevrolet	1	362	9.7%
Dodge	2	250	6.7%
Ford	3	432	11.6%
GMC	4	139	3.7%
Honda	5	577	15.5%
Hyundai	6	183	4.9%
Jeep	7	146	3.9%
Mazda	8	173	4.7%
Nissan	9	449	12.1%
Toyota	10	1,003	27.0%
Total		3,714	100.0%

captured by duration (in seconds, LMD[11]: −3,403.7) than by the number of ads (LMD: −3,443.2; Log Bayes Factor[12]: 39.5). Due to high colinearity between duration and number of ads, including both measures in the model does not improve the fit (LMD: −3,436.1; LBF[12]: 32.4). Lastly, we note that the best model which includes banner ads as a covariate is selected over the best model which does not include banner ads as a covariate (LMD: −3,434.2, LBF[12]: 30.5).

4.2 Model validation with a holdout sample

For each individual we keep two observations for the holdout test. We measure predictive fit as the absolute mean deviation (MAD) between the observed choices and the choice probabilities given by our respective models. The choice probabilities are calculated using the mean of the posterior distribution of each individual's parameter estimates. The holdout results confirm the in-sample findings above: the best fitting model is a two-component Mixture (MAD = 0.244, see Table 6) outperforming a three-component Mixture (MAD = 0.293) and a standard continuous heterogeneity (or one-component) model (MAD = 0.332).

4.3 Parameter estimates

We find that expectations (based on previous research) are verified in the results (see Table 7). First, current and previous session make exposures are good predictors for the choice of the next make-page in both segments (like last-brand purchased in scanner research). Note the 3:1 ratio of current vs. last session make exposures across segments. Consumers' current session state dependence is found

[11] Based on the best fitting two-component Mixture model with logarithmic functional form (f) and $K^{make} = 2$ and $K^{ad} = 1$.

[12] A Log Bayes Factor (LBF) above 8 indicates very strong evidence that the focal model is preferred over the alternative model. LBFs shown here are calculated with respect to the two Component Mixture model (i.e., the focal model) with LMD of −3,403.7.

Table 4 Sample properties

Make	Make page views	
	Number	Percent
Chevrolet	745	9.7%
Dodge	517	6.7%
Ford	1,142	14.8%
GMC	278	3.6%
Honda	1,079	14.0%
Hyundai	217	2.8%
Jeep	222	2.9%
Mazda	296	3.9%
Nissan	1,028	13.4%
Toyota	2,166	28.2%
Total	7,690	100.0%
Make ID	Banner Ad Views	
	Number	Percent
Ford	1,649	26.8%
GMC	2,137	34.7%
Toyota	2,369	38.5%
Total	6,155	100.0%

to be three times as important as information from the last session. Second, banner ad exposure, in contrast, is only effective during the *current* session. Ad exposure from *previous* sessions does not influence the choice of the next make-page. This is consistent with previous research which found that banner ads had decaying effectiveness in the automobile category (Havlena and Graham 2004). Third, we find that consumers react differently to ad exposure. Segment 1 has a positive response to advertising, whereas segment 2 does not respond (see Table 7).

Does the implied segmentation allow us to understand differences in response to make and ad exposure? As discussed in more detail below, we find that browsing behavior differs across segments in terms of page-views and make-views. This gives rise to a potential explanation of (make) state-dependence and advertising response.

Table 5 Model comparison

Model[a]	Log marginal density	Log bayes factor[b]
One component mixture	−5,196.6	1,792.9
Two component mixture	−3,403.7	–
Three component mixture	−3,451.2	47.5

[a] We use $K^{make} = 2$ and $K^{ad} = 1$; the effects of banner ads is measured in duration (seconds) and past exposures to makes and banner ads are modeled with a logarithmic formulation, i.e., they have diminishing returns. This specification gives the best fit

[b] A Log Bayes Factor above 8 is very strong evidence that one model is superior in fit to another. Log Bayes Factors shown here are calculated with respect to the Two Component Mixture model.

Table 6 Holdout performance

Model	Mean absolute deviation (MAD)
One component mixture	.332
Two component mixture	.244
Three component mixture	.293

We discuss the browsing behavior for each segment in detail and show how differences in response could be related to differences in browsing behavior. Again, clickstream data does not allow us to identify the individuals, so we cannot, for example, investigate whether the segments differ based on demographics.

A Mixture approach allows us to calculate an individual's probability of belonging to a segment based on the draws of the segment-indicators Z. The majority of consumers (219 out of 250) can be classified with a probability greater than 0.8 (see Table 8). Of the remaining 31 consumers, we find that 23 consumers can be classified with a probability exceeding 0.6. The residual 8 can only be classified with a probability between 0.5 and 0.6. We classify 134 (53.6%) consumers as belonging to segment 1, or the ad-responsive segment, and 116 (46.4%) consumers as belonging to segment 2, or the non-ad-responsive segment. Using this classification, we investigate differences in users' make-page choices in lieu of demographics.

Starting with segment 1, we find that the average number of page views is significantly above the sample average (141.3, compared to 104.3 for the total sample, see Table 9) and significantly above the average 61.7 page-views of segment 2. We posit that consumers in segment 1 behave in a less focused manner (i.e., more

Table 7 Parameter estimates for the two component mixture model

	Component 1		Component 2	
	Mean	95% coverage interval	Mean	95% coverage interval
Make current session	**1.274**	**(1.064, 1.466)**	**8.594**	**(7.834, 9.329)**
Make last session	**0.394**	**(0.173, 0.592)**	**3.086**	**(2.575, 3.629)**
Ads current session	**0.178**	**(0.151, 0.203)**	−0.085	(−0.848, 0.404)
Intercepts				
Dodge	−0.097	(−0.732, 0.501)	0.185	(−0.542, 0.939)
Ford	**0.608**	**(0.124, 1.362)**	**−2.645**	**(−4.408, −1.015)**
GMC	**−1.185**	**(−1.590, −0.731)**	−1.049	(−2.749, 0.416)
Honda	**1.479**	**(0.613, 2.362)**	−0.368	(−1.435, 0.716)
Hyundai	**−1.461**	**(−2.062, −0.849)**	**−3.435**	**(−5.210, −1.519)**
Jeep	**−1.865**	**(−2.572, −1.086)**	**−2.777**	**(−4.492, −1.258)**
Mazda	−0.557	(−1.275, 0.209)	**−1.297**	**(−2.367, −0.331)**
Nissan	**1.411**	**(0.698, 2.155)**	−0.690	(−1.744, 0.308)
Toyota	**2.210**	**(1.561, 2.970)**	**1.257**	**(0.394, 2.171)**

Significant estimates are in boldface

Table 8 Segment assignment performance

Probability of individual belonging to segment 1[a]	Segment 1	Segment 2	Total
0–10%	0	84	84
11–20%	0	18	18
21–30%	0	8	8
31–40%	0	4	4
41–50%	0	2	2
51–60%	6	0	6
61–70%	4	0	4
71–80%	7	0	7
81–90%	22	0	22
91–100%	95	0	95
Total	134	116	250

[a] Probability of belonging to Segment 2 = 1-Prob(Segment 1)

of a browsing mode) when compared to those in segment 2. Consistent with this notion, previous make exposures are less predictive of the consumers' make-page choices for segment 1 than for segment 2. Thus, compared to segment 2, the consumers in segment 1 should be more receptive to banner advertising. In line with expectations, consumers in segment 1 react positively to banner advertising–ad exposure leads to more subsequent search for the featured make.

We now turn our analysis to the non-ad-responsive segment–segment 2. Consumers in segment 2 have a lower number of mean page views (61.7, compared to 104.3 for the total sample, see Table 9). This could indicate more focused search by consumers who already know what they are interested in (Moe 2003; Montgomery et al. 2004). As make-page views and ad views are highly correlated with page views, consumers in this segment also have less exposure to makes and ads than segment 1. In segment 2, we also find that the makes previously viewed are a very strong predictor of the next make-page choice. This could be interpreted as

Table 9 Segment characteristics

	Segment	
	1	2
Size (in%)	53.6%	46.4%
Mean page views	141.3	61.7
Mean make views	43.1	16.6
Mean make choices	24.3	8.8
Mean ads duration (in seconds)	521.5	273.2
Mean make page views per page view	0.30	0.27
Mean make choices per page view	0.17	0.14
Mean ads duration (in seconds) per page view	3.69	4.43

consumers being in a goal-directed mode. As Danaher and Mullarkey (2003) note, consumers in a goal-directed mode are much less likely to recall and recognize banner ads than consumers who are in a browsing mode. This is consistent with what we find–in segment 2 consumers' make-page choices are not influenced by banner ad exposure.

4.4 Alternative explanations

In this section, we discuss possible alternative explanations for our findings. For example, one could imagine that consumers in segment 2 (non-ad responsive) simply look at very few makes compared to the other segment (i.e., they make very few make-page choices). As a result, we might obtain an insignificant effect due to very few observations. On the other hand, the ratio of make-pages to total page-views is stable across segments (see Table 9). Thus, consumers do not differ across segments in their propensity to look at make- and non-make-pages. The same holds true when we look at the numbers of make choices and ad duration relative to the number of total page views. Again, we find very little difference among the segments (see Table 9).

We observed banners ads for only three makes (Ford, GMC, and Toyota). A second alternative explanation is that our results may be distorted by selection bias. In other words, could ad responsiveness (segment 1) simply be driven by consumers who look mostly at the Ford, GMC, and Toyota makes and submitted leads for one of these makes? If this were the case, we should find most consumers who submitted leads for a non-advertised make to be in the other, non-ad responsive, segments (i.e., segment 2). To rule out selection bias as a possible explanation, we examined the distribution of leads by make in the segments versus the entire sample. Table 10 shows that the distribution of the leads in the two segments closely resembles the distribution of the leads in the total sample. Thus, the selection bias described above seems an unlikely explanation for our results.

Table 10 Lead distribution across segments

Make		Segment 1	Segment 2	Total
1	Chevrolet	9.0%	7.1%	7.6%
2	Dodge	6.0%	4.3%	5.2%
3	**Ford**	**16.2%**	**11.1%**	**12.0%**
4	**GMC**	**2.7%**	**7.4%**	**4.0%**
5	Honda	12.6%	13.0%	14.0%
6	Hyundai	3.1%	2.9%	2.8%
7	Jeep	4.0%	5.4%	4.0%
8	Mazda	3.1%	4.5%	4.0%
9	Nissan	11.2%	13.0%	14.8%
10	**Toyota**	**32.1%**	**31.3%**	**31.6%**

Advertised makes appear in boldface

4.5 Banner ad elasticities

Quantifying the effects of advertising has long been a focus of marketing research. A consistent finding of this body of work is that the advertising elasticity with respect to sales is in the range of 0.2 (e.g., Tellis 2004). Compared to traditional advertising, the effects of banner ads on sales are hard to track and, we speculate, are very small in the case of durables. Our model allows us to shed new light on the effectiveness of banner ads on a different dimension. We do not investigate their effect on sales, but focus on changes in behavior that can be identified from clickstream data. We use the parameter estimates of our model to calculate the elasticities of make-page choice with respect to changes in banner ad exposure.

We calculated the advertising make-page elasticities for segment 1 across the three different makes which placed premium display ads on the site (see Table 11). We find that the elasticities for segment 1 are not far from the average advertising elasticity of 0.2. In our data, the advertising elasticity also differs with respect to make: GMC has the highest elasticity with 0.17, followed by Ford with 0.15 and Toyota at 0.11. We caution that these are not sales or volume elasticities with respect to advertising, but pertain to the influence of the ads on website page choice and for consumers in the most responsive segment.

4.6 Managerial implications

Ad testing One area in which our model is directly applicable is ad testing. Traditionally, different ad copies (or even campaigns) are run using test and control conditions and then the performance of these campaigns is evaluated using survey methods. This can be both expensive and time consuming. Our model allows firms to assess one dimension of the effectiveness of Internet display ads by analyzing clickstream data. Ford, for example, could set up a test for a banner ad for a Ford Explorer. Using clickstream data, the effectiveness of the ad, versus the appropriate control, could be assessed without the need to conduct a survey.

Firms could also use the model to continuously monitor the effectiveness of a given ad or campaign. By comparing model results over time, managers might be able to detect the onset of ad wear-out and rotate different ad copies based on these findings. This could provide new input into procedures for refreshing ads and optimizing ad rotation.

Pricing ads Our findings are also important for website managers. This is particularly true when the site is dependent upon ad revenue. We find that a banner ad leads to more exposure for the advertised make by positively influencing the probability that the advertised make is viewed by a segment of consumers.

Table 11 Advertising elasticities

Make	Elasticity
Ford	0.15
GMC	0.17
Toyota	0.11

Management can use our findings to help sell more advertising and potentially adjust pricing policies (e.g., different sections of the website might be priced differently depending upon how effective they are). Pricing or valuing banner ads depending on their display location within the site is an interesting topic for future research.

Targeting ads Another potential future application of our findings and approach could be better targeting of ads to the web site's users. A similar model could be used to generate a segmentation of ad responsiveness after, for example, the first 30 page views. Based on the results, consumers could be better targeted by giving the positive response segments more ads and non-responsive segment fewer ads. This seems technically feasible, but requires significant investment and a move to a truly interactive website. Our findings could be helpful in setting up the segmentation algorithm.

5 Conclusion

We propose a new approach to investigating the effectiveness of banner ads that differs from previously used measures such as impressions served, click-through rates, brand attitudes (e.g., recall), and final outcome measures (e.g., purchase). Using Internet clickstream data, we can address an important dimension of the effect of banner ads that has not yet been explored: do banner ads have an immediate influence on brand search behavior? More specifically, do banner ads alter the make-page choices of the consumer subsequent to exposure? Our approach could be particularly valuable for durable goods because advertising exposure is often further removed in time from purchase than for non-durables. If this turns out to make the effect of banner advertising on durable good purchase difficult to determine, advertisers and web sites would still be able to use our approach to shed light on how banner ads alter more immediate forms of consumer behavior.

Using clickstream data from an automotive website, we model the brand-specific page view choices a consumer makes during a session. While on the site, consumers are exposed to banner advertising for different car makers. Our study focuses on how these ad exposures do or do not affect a user's subsequent sequence of make-page choices–each of which provides website content for the requested makes. We model these make-page choices as a sequence of logit choices that dynamically use sequencing and duration information. Thus, a previous choice influences a current choice in our framework. We estimate our model using a Bayesian Mixture approach to overcome the problems that continuous heterogeneity models have with multimodal distributions.

A two-component Mixture model with banner ad exposure as a covariate fits the data best in sample as well as in hold-out. In this two-segment model, the 54% of consumers in segment 1 react positively to banner advertising while those in segment 2 do not. Using a standard (unimodal) Bayesian set-up to account for heterogeneity would have led to us to *incorrectly* conclude that banner advertising does not affect browsing behavior (advertising is not significant in the unimodal set-up). A Bayesian Mixture model allows us to incorporate the multimodality apparent in the response to banner advertising and correctly recover the individual-level parameters. Based on these parameters we calculate an average elasticity for make-page choice with respect to banner advertising exposure of about 0.15, which is in the range of the average of 0.2 reported in most advertising studies. Because we do not study the effect of banner advertising on final purchase, our findings are limited to its effect on the make-page choice decisions in the user's browsing history.

In the model, current and previous exposure to make pages reveals make preferences, controls for state-dependence, and allows good prediction of the next make-page choice. As expected, current make-page exposure has a stronger effect than that from the previous session; for both segments the ratio of current vs. previous exposure is roughly 3:1. The segments each have different levels for these effects with segment 2, the non-ad responsive segment, having the stronger make effects.

Our findings are important from a managerial perspective because we are able to show that banner ads do affect the subsequent behavior of some users on the web site. This effect cannot be captured by previously employed measures such as click-through rates or outcome measures such as brand choice or purchase incidence. Our findings help shed light on the puzzle of banner advertising: academically mixed evidence for the effectiveness of banner ads juxtaposed with growing spending by advertisers.

Our approach can be used to test ads without expensive procedures such as surveys or eye-tracking. Firms can compare the effectiveness of different ads, for example in terms of ad copy or placement on the site, by modeling brand-specific page-view choices using clickstream data. Our approach allows "online" monitoring of this aspect of campaign performance and enables firms to rapidly detect wearout of individual ads. This could enable firms to better optimize ad rotation and improve advertising effectiveness. At the same time, our finding that banner ads "work" could strengthen the position of the web site owner versus the advertiser when it comes to selling and pricing advertising.

Our study also has a number of limitations. First, we did not connect the observed behavior and the user's final make choice, submitted in the form of a lead, within an integrated model. This is because submitting a lead for a new car is far from the same as making a purchase; the conversion rate from submitted leads to purchase was about 10%. Thus, our dataset is too small to model purchase as the dependent variable. Second, we have not been able to fully explore the reasons why some consumers are responsive to advertising and some are not. Additional demographic information might be useful to further investigate this behavior. A final limitation is that our data contain information only for premium display ads and do not include non-premium or remnant banner ads. While we checked for possible selection bias and did not find evidence for it, future clickstream studies of banner advertising would benefit from full information regarding advertising placement and exposure.

Acknowledgement The authors wish to thank a collaborating firm for providing the data used in this study.

Appendix: Sampler

Based on a set of staring values $\left[\theta, Z, V, \widehat{Q}\right]$ repeat the following steps:

Step 1: Propose θ_i^{new} and use a Metropolis-Hastings step to accept/reject θ_i^{new} for all i=1,…,n based on the logit-likelihood defined in (3).
Step 2: For s=1,…,S draw new μ_s and $\mathrm{V_s}$ using Gibbs sampling

- $\mu_s \Big| \theta, Z, V_s, \mu_0, V_0 \sim N\left(\widetilde{\mu}_s, \widetilde{V}_s\right)$, where $\widetilde{V}_s = \left[\left(\sum_{i=1}^{n} I(Z_i = s)\right) V_s^{-1} + V_0^{-1}\right]^{-1}$ and $\widetilde{\mu}_s = \widetilde{V}_s \left[\left(\sum_{i=1}^{n} I(Z_i = s)\theta_i\right) V_s^{-1} + V_0 \mu_0\right]$, where $I(Z_i$=$s)$ is the indicator function (1 if Z_i=s and 0 otherwise).
- $V_s^{-1} | \theta, Z, \mu_s \sim Wishart\left(v_1 + \sum_{i=1}^{n} I(Z_i{=}s), \left[v_2 + \sum_{i=1}^{n} I(Z_i = s)(\theta_i - \mu_i)(\theta_i - \mu_s)'\right]^{-1}\right)$, where $I(Z_i$=$s)$ is the indicator function (1 if Z_i=s and 0 otherwise).

Step 3: Draw new weights π using Gibbs sampling

- $\pi | Z, \rho \sim Dir[(\widetilde{\rho}_1 \ldots \widetilde{\rho}_S)]$, where $\widetilde{\rho}_s = \rho_s + \sum_{i=1}^{n} I(Z_i = s)$ with prior ρ=(1,…,1).

Step 4: For i=1,…,n draw new Z_i using Gibbs sampling

- $Z_i | \theta, \mu, V, \pi \sim multinomial\left(1, \left[\frac{\pi_1 \phi(\theta_i | \mu_1, V)_1}{\sum_{s=1}^{S} \pi_s \phi(\theta_i | \mu_s, V)_s}, \ldots, \frac{\pi_S \phi(\theta_i | \mu_S, V_S)}{\sum_{s=1}^{S} \pi_s \phi(\theta_i | \mu_s, V)_s}\right]\right)$, where ϕ is the probability density function of the Gaussian distribution.

Step 5: Choose a permutation ξ_t^+ and relabel draws according to ξ_t^+

$$\xi_t^+ = \arg\min_{\xi_t} \sum_{i=1}^{n} \sum_{s=1}^{S} p_{is}(\xi_t(\mu^t, V^t)) \log\left[\frac{p_{is}(\xi_t(\mu^t, V^t))}{\widehat{q}_{is}^{t-1}}\right].$$

Step 6: Set

$$\widehat{Q}^t = \frac{t\widehat{Q}^{t-1} + P\left(\xi_t^*(\mu^t, V^t)\right)}{t+1}.$$

References

Allenby, G. M., Arora, N., & Ginter, J. L. (1998). On the heterogeneity of demand. *Journal of Marketing Research, 35*, 384–398.

Avenue A/Razorfish (2006). *In business week "The Future of Technology,"* No. 56, March 27.

Andrews, R. L., Ainslie, A., & Currim, I. S. (2002). An empirical comparison of logit choice models with discrete vs. continuous heterogeneity. *Journal of Marketing Research, 39*, 479–487.

Bass, F., Bruce, N., Majundar, S., & Murthi, B. P. S. (2007). A dynamic Bayesian model of the advertising-sales relationship. *Marketing Science, 26*(2), 179–195.

Bucklin, R. E., & Sismeiro, C. (2003). A model of web site browsing behavior estimated on clickstream data. *Journal of Marketing Research, 40*(2), 249–267.
Business Week (2006). "*The future of technology*," No. 56, March 27.
Business Week (2007). "*So many ads, so few clicks*," No. 46, November 12.
Chatterjee, P., Hoffman, D. L., & Novak, T. P. (2003). Modeling the clickstream: Implications for web-based advertising efforts. *Marketing Science, 22*(4), 520–541.
Celeux, G., Chauveau, D., & Diebolt, J. (1996). Stochastic versions of the Em algorithm: An experimental study in the mixture case. *Journal of Statistical Computation and Simulation, 55*, 287–314.
Celeux, G. (1998). "Bayesian inference for mixtures: The label-switching problem." In *COMPSTAT 98*, 227–232, Heidelberg, Physica.
Cho, C.-H., & Cheon, H. J. (2004). Why do people avoid advertising on the internet? *Journal of Advertising, 33*(4), 89–97.
ComScore (2008). "*General Motors Ranks as the heaviest online advertiser among auto manufacturers in January, according to comScore Ad Metrix*." Press Release, March 27.
Danaher, P. J., & Mullarkey, G. W. (2003). Factors affecting online advertising recall: A study of students. *Journal of Advertising Research, 43*(3), 252–267.
Diebolt, J., & Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society: Series B, 56*(2), 363–375.
DoubleClick (2003). "*DoubleClick 2002 full-year ad serving trends*," www.doubleclick.net, accessed January 2003.
Dreze, X., & Husshérr, F.-X. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing, 17*(4), 8–23.
Dube, J.-P., Hitsch, G., & Rossi, P. E. (2009). Do switching costs make markets less competitive. *Journal of Marketing Research, 46*, 435–445.
Dube, J.-P., Hitsch, G., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics, 41*(3), 417–445.
Geweke, J. F., Keane, M., & Runkle, D. (1997). Statistical inference in the multinomial multiperiod probit model. *Journal of Econometrics, 80*, 125–166.
Goldfarb, A. (2002). "Analyzing website choice using clickstream data." In: M. Baye (Ed.), *Advances in applied microeconomics v. 11: The economics of the internet."* Elsevier Science Ltd., 209–230.
Hanssens, D. M., Parsons, L. J., & Schultz, R. L. (1990). *Market response models: Econometric and time series analysis*. Boston: Kluwer Academic Publishers.
Havlena, W., & Graham, J. (2004). Decay effects in online advertising: Quantifying the impact of time since last exposure on branding. *Journal of Advertising Research, 44*, 327–332.
Hubermann, B. A., Priolli, P. L. T., Pitkow, J. E., & Lukose, R. M. (1998). Strong regularities in world wide web surfing. *Science, 280*(3), 95–97.
Johnson, E. J., Moe, W., Fader, P. S., Bellman, S., & Lohse, G. L. (2004). On the depth and the dynamics of online search behavior. *Management Science, 50*(3), 299–308.
Klein, L. R. (1998). Evaluating the potential of interactive media through a new lens: Search versus experience goods. *Journal of Business Research, 41*(3), 195–203.
Lodish, L., Abraham, M., Kalmensen, S., Livelsberger, J., Lubetikin, B., Richardson, B., et al. (1995). How TV advertising works: A meta analysis of 389 real-world split-cable TV advertising experiments. *Journal of Marketing Research, 32*(2), 125–139.
Manchanda, P., Dube, J.-P., Goh, K. Y., & Chintagunta, P. K. (2006). The effects of banner advertising on internet purchasing. *Journal of Marketing Research, 43*(1), 98–108.
McCulloch, R., & Rossi, P. E. (1996). Bayesian analysis of the multinomial probit model. In R. Mariano, Weeks, T. Schuermann (Eds.), *Simulation-based inference in econometrics*. Cambridge: Cambridge University Press.
Moe, W. (2003). Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of Consumer Psychology, 13*(1&2), 29–39.
Moe, W., & Fader, P. S. (2004). Dynamic conversion behavior at E-commerce sites. *Management Science, 50*(3), 326–335.
Moe, W. (2006a). An empirical two-stage choice model with varying decision rules applied to internet clickstream data. *Journal of Marketing Research, 43*(4), 680–692.
Moe, W. (2006b). A field experiment to assess the interruption effect of pop-up promotions. *Journal of Interactive Marketing, 20*(1), 34–44.
Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science, 23*(4), 579–595.

Moore, R. S., Stammerjohan, C. A., & Coulter, R. A. (2005). Banner advertiser-web site congruity and color effects on attention and attitudes. *Journal of Advertising, 34*(2), 71–84.

Naik, P. A., Mantrala, M., & Sawyer, A. (1998). Planning media schedules in the presence of dynamic advertising quality. *Marketing Science, 17*(3), 214–235.

Newstead, K., Taylor, J., Kennedy, R., & Sharp, B. (2009). The total long-term sales effects of advertising: Lessons from single source. *Journal of Advertising Research, 49*(2), 207–210.

Novak, T. P. & Hoffman, D. L. (2000). "Advertising and pricing models for the web." In: *Internet publishing and beyond: The economics of digital information and intellectual property*, Cambridge: MIT Press.

PricewaterhouseCoopers (2010). *IAB Internet Advertising Revenue Report*, http://www.iab.net/media/file/resources_adrevenue_pdf_IAB_PwC_2010_Final.pdf.

Richardson, S., & Green, P. J. (1997). On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society: Series B, 59*, 702–731.

Rubinson, J. (2009). Empirical evidence of TV advertising effectiveness. *Journal of Advertising Research, 49*(2), 220–226.

Sismeiro, C., & Bucklin, R. E. (2004). Modeling purchase behavior at an E-commerce web site: A task completion approach. *Journal of Marketing Research, 41*(3), 306–323.

Stephens, M. (2000). Dealing with label switching in mixture models. *Journal of the Royal Statistical Society: Series B, 62*(4), 795–809.

Tellis, G. J. (2004). *Effective advertising: How, when, and why advertising works*. Thousand Oaks: Sage Publications.

Terui, N., Ban, M., Allenby, G. M. (2010)."The effect of media advertising on brand consideration and choice." *Marketing Science, 30*(1), 74–91.

ThinkEquity Partners (2007). *Online advertising 2.0*. September 15, San Francisco.