Competitive Analytics of Multi-channel Advertising and Consumer Inertia

(Yiyi Li
UT Dallas
yxll10020@utdallas.edu)

(Ying Xie
UT Dallas
ying.xie@utdallas.edu)

(Eric Zheng
UT Dallas
zxz062000@utdallas.edu)

Abstract

The spurs and multiplicity of Internet advertising have made multi-channel attribution an immediate challenge to marketing practitioners. We propose an integrated individual-level choice model that considers three stages of a consumer’s purchase funnel – awareness, alternative evaluation and purchase - across all competitors to analyze the effects of touches on (1) consumers’ choice of entry site, (2) their subsequent search decisions about other websites in the awareness set, and (3) their subsequent purchases at one of the searched websites.

Keywords

Multichannel marketing, purchase funnel, advertising response.

Introduction

Advancement of digital technologies has enabled firms to reach customers through a variety of ad channels such as search engine, email, referral sites and display ads. As a result, spending on digital advertising has reached $42.3 billion in 2013 in the U.S. (emarketer.com, 2013). Ad attribution, the analytics ads effectiveness across multi-channels and multi-touchpoints (MCMT), has become an immediate challenge to marketing practitioners (Daleasandro et al. 2012).

However, most companies evaluate the effectiveness of different Internet advertising formats based on the sequence of the ads in the conversion path under various heuristics rules. Examples include the single source attribution models such as the last interaction model and the first interaction model, and the fractional attribution models such as the linear model, the time decay model and the position based model. These attribution models commonly weight the advertisements in an ad-hoc manner, e.g. simply assigning weights based on their appearance sequence in the purchase paths. These heuristic attribution rules have been criticized to be largely inconsistent and inconclusive, and they only consider the paths resulting in conversions while disregarding the majority of paths that do not, leading to biased insights (Li and Kannan 2014). For these reasons, the industry calls for a more scientific measure that is data-driven (Econsultancy 2012, Shao et al. 2011).

Although the academic literature in this area has been sparse, an increasing number of algorithmic attribution models have been proposed in recent years (Li and Kannan 2014, Abhishek et al. 2014, Xu et al. 2014). These models predominantly focus on consumers’ interaction with the focal website only, disregarding the impact of competitive actions on the conversion probability, a critical determinant of ad effectiveness. As is common in today’s hyper-competitive online retailing world, customers are likely to be exposed to various forms of digital ads across all the competing firms in a particular product category.
Failing to account for customers’ across-competitor interactions may yield biased estimate of ad effectiveness, as we demonstrate in this paper.

This study aims to fill the gap by providing a new ad performance measure in a competitive online shopping environment. We achieve our objective by modeling consumer’s multinomial choice decisions across all competing websites in contrast to a binary choice between conversion and no conversion with respect to a focal site only.

Further, we examine the entire funnel of a customer’s purchase path. The literature suggests that the effectiveness of various digital ads depends on a consumer’s readiness stage in his/her online shopping journey (Abhishek et al. 2014). In general, a consumer goes through three stages in a purchase session: 1) awareness stage, in which she perceives a need and form a consideration set; 2) alternative evaluation stage, in which she visits one or several stores to search for relevant information; 3) the final purchase stage. In the online purchasing funnel, Ratchford (2008) documents that the amount of online search is limited despite greatly lowered search costs. For example, the average number of unique sites searched by each household in the air-travel industry is around 1.8 and 42% of the travel shoppers visit only one website (Johnson et al. 2004). This phenomenon indicates that being the entry site (i.e., the first website visited by the consumer during a purchase journey) is vital. From the marketer’s perspective, it is ideal for a consumer to discontinue her search process and convert at the entry site. Our model adopts the funnel view of consumer’s online purchase path, and specifically models customers’ entry decisions.

Lastly, consumers’ loyalty plays a critical role in understanding the funnel. Repeat buys account for over half of e-tailers’ sales (CyberAtlas, 2002). Consumer’s search and purchase decisions are affected not only by various forms of advertising but also by past interactions with the websites. Therefore, it is essential for our model to account for state dependence in order to capture the dynamics of consumers’ online behaviors. In particular, we need to identify 1) whether the observed inertia is due to structural state dependence or simply induced by heterogeneity in consumer preference, and 2) if it is the former, what the source of such inertia is.

In this paper, we propose an integrated three-stage choice model that considers individual customer’s search and purchase decisions across all competitors to analyze the effects of ad clicks on (1) consumer search, where we specifically examine each customer’s choice of entry site and visit decision for other websites they are aware of, and (2) their subsequent purchases at one of the searched sites. This model is estimated using a unique disaggregated individual level panel dataset that records consumer interactions with all relevant websites in the online air ticket booking industry through various online advertising channels including search engine, email, display ads and referrals (non-paid links from other websites) along a conversion path. Our result shows that different types of digital ads affect consumers differently based on the stages in their purchase funnel. Email and referrals affect visit decision positively, but have limited impact on conversion probability, while search engine and display ads have a positive impact on both decisions. Our proposed model suggests that the direct effect of advertising on purchase is smaller in size in general compared with the benchmark model that focuses solely on purchase decision. Secondly, we find that inertia of search and purchase is robust to flexible controls for preference heterogeneity, and we quantify the long-term margins of advertising contacts. Our model better predicts individual consumer’s online behavior based on their past behavioral data compared with the model that does not consider consumer dynamics. Third and more importantly, with the current model, we are able to compute the own- and cross-marginal impact of ad channels for each website and find that the effectiveness of ad channels varies across websites.

**Empirical Setting**

Our data comes from an online panel provided by a leading Internet analytics company. The raw data includes detailed online browsing and transaction data from a group of Internet users chosen randomly by the data provider for the entire year of 2009. Each user’s online activities are recorded by the proxy servers to form a click-stream dataset that includes the user id, domain name, date, time, duration, number of pages viewed, redirecting website, and transaction related information including product description, category, price, quantity and basket total if a purchase is observed. This dataset is unique as it contains consumer’s every interaction with all relevant websites in a purchase session. In addition, we are able to infer the advertising channel used by the consumer from the redirecting website. A third
advantage of our dataset is that it records each consumer’s complete browsing and purchasing history with all websites, so we are able to capture the dynamics in consumer decisions.

We focus on the air-travel industry. Our sample comprises of 1697 purchase sessions made by 1002 panelists. We define a purchase session as all visits to relevant websites in the industry within 7 days prior to a purchase, following De Los Santos et al. (2014). A website is considered relevant if there is at least one air-travel purchase occurred at the website during our data time period. We exclude Southwest.com from our analysis because it is distinct from the rest of major travel sites. Table 1 presents the complete list of websites included in our analysis. These websites are grouped into: 1) online direct channel of airline companies, 2) online agents that sell tickets of multiple airlines. To avoid the curse of dimensionality in estimation, we focus on the competition among online agents, and group all airline companies together under the label “direct channel”. We also group americanexpress-travel.com, sabresonicweb.com and wweet1.com together and label them as “small agents” since the number of visits and transactions of these three agents are significantly smaller than the others.

We group all the redirecting websites into four channels: search engine, email, display ads, and referrals (advertising on third-party websites that the firms do not pay for). The null option is direct visit by typing the website address. The top five redirecting websites in each channel are shown in Table 2.

Table 1: List of Websites in Air Ticket Category

<table>
<thead>
<tr>
<th>Code</th>
<th>Websites</th>
<th># of Searches</th>
<th># of Purchases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Airline direct</td>
<td>1306 (35.70%)</td>
<td>1027 (60.52%)</td>
</tr>
<tr>
<td>2</td>
<td>cheaptickets.com</td>
<td>259 (7.08%)</td>
<td>86 (5.07%)</td>
</tr>
<tr>
<td>3</td>
<td>expedia.com</td>
<td>604 (16.51%)</td>
<td>201 (11.84%)</td>
</tr>
<tr>
<td>4</td>
<td>hotwire.com</td>
<td>180 (4.92%)</td>
<td>29 (1.71%)</td>
</tr>
<tr>
<td>5</td>
<td>orbitz.com</td>
<td>451 (12.33%)</td>
<td>154 (9.07%)</td>
</tr>
<tr>
<td>6</td>
<td>priceline.com</td>
<td>403 (11.02%)</td>
<td>86 (5.07%)</td>
</tr>
<tr>
<td>7</td>
<td>Small agents</td>
<td>53 (1.45%)</td>
<td>9 (0.53%)</td>
</tr>
<tr>
<td>8</td>
<td>travelocity.com</td>
<td>402 (10.99%)</td>
<td>105 (6.19%)</td>
</tr>
</tbody>
</table>

Table 2: Top Redirecting Websites in Each Advertising Channel

<table>
<thead>
<tr>
<th>Search Engine</th>
<th>Email</th>
<th>Display Ads</th>
<th>Referrals</th>
</tr>
</thead>
<tbody>
<tr>
<td>google.com</td>
<td>live.com</td>
<td>hotels.com</td>
<td>kayak.com</td>
</tr>
<tr>
<td>yahoo.com</td>
<td>comcast.net</td>
<td>clxtrkr.com</td>
<td>lowfares.com</td>
</tr>
<tr>
<td>aol.com</td>
<td>msn.com</td>
<td>conduit.com</td>
<td>bookingbuddy.com</td>
</tr>
<tr>
<td>bing.com</td>
<td>rr.com</td>
<td>alot.com</td>
<td>travelzoo.com</td>
</tr>
<tr>
<td>ask.com</td>
<td>att.net</td>
<td>port-columbus.com</td>
<td>cheapflights.com</td>
</tr>
</tbody>
</table>

Table 3: Channel Usage for One Purchase Session

<table>
<thead>
<tr>
<th>Code</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>search engine</td>
<td></td>
<td>email</td>
<td></td>
<td>display</td>
<td></td>
<td>referrals</td>
<td>direct visit</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.80 (1.61)</td>
<td>0.44 (0.82)</td>
<td>0.40 (0.84)</td>
<td>0.42 (1.00)</td>
<td>0.31 (0.75)</td>
<td>0.30 (0.74)</td>
<td>0.01 (0.11)</td>
<td>0.39 (0.79)</td>
<td>0.51 (1.15)</td>
</tr>
<tr>
<td></td>
<td>0.06 (0.35)</td>
<td>0.04 (0.24)</td>
<td>0.03 (0.22)</td>
<td>0.03 (0.24)</td>
<td>0.03 (0.33)</td>
<td>0.02 (0.18)</td>
<td>0.00 (0.00)</td>
<td>0.02 (0.17)</td>
<td>0.04 (0.28)</td>
</tr>
<tr>
<td></td>
<td>0.43 (1.40)</td>
<td>0.02 (0.16)</td>
<td>0.14 (0.45)</td>
<td>0.03 (0.20)</td>
<td>0.04 (0.20)</td>
<td>0.02 (0.16)</td>
<td>0.00 (0.00)</td>
<td>0.02 (0.17)</td>
<td>0.17 (0.53)</td>
</tr>
<tr>
<td></td>
<td>0.07 (0.36)</td>
<td>0.09 (0.34)</td>
<td>0.09 (0.36)</td>
<td>0.13 (0.36)</td>
<td>0.17 (0.55)</td>
<td>0.21 (0.60)</td>
<td>0.11 (0.37)</td>
<td>0.08 (0.32)</td>
<td>0.11 (0.42)</td>
</tr>
<tr>
<td></td>
<td>2.14 (2.87)</td>
<td>1.10 (1.22)</td>
<td>1.34 (1.99)</td>
<td>1.08 (1.31)</td>
<td>1.23 (1.37)</td>
<td>1.19 (1.18)</td>
<td>0.63 (0.79)</td>
<td>1.15 (1.64)</td>
<td>1.50 (2.12)</td>
</tr>
</tbody>
</table>
We observe ad click-throughs but not ad exposures, and so we aggregate the number of click-throughs from each ad channel to represent the information stock for each website. The average number of click-throughs per purchase session for each ad channel/website combination is presented in Table 3 (along with the standard deviation in parenthesis). On average, search engine is the most widely used redirecting channel, followed by display ads.

The Model

Our model adopts the purchase funnel view in the context of online purchases of high-involvement products in a competitive environment. The purchase funnel is grounded in the information processing theory that describes the way consumers make their purchase decisions, from perceiving a need all the way to post-purchase behaviors (Bettman et al., 1998). Figure 1 illustrates how a consumer moves through the three stages during a purchase session.

![Figure 1: A Funnel View of Online Purchase](image)

Different ad format affects consumers differently based on its stage in a purchase funnel. In order to capture the effectiveness of different ad formats at different stages, we propose a three-level model which explicitly accounts for the formation of awareness set, the visit decision and the purchase decision in a competitive environment. The details of each stage are discussed as follow.

**Stage 1: Awareness**

At this stage, the consumer conceives a purchase need and retrieves a set of websites that may satisfy the need from memory. Such awareness of these websites due to past ad exposures, interactions with the websites or word-of-mouth. Such awareness affects subsequent decisions such as which websites to visit and where to make the final purchase. Traditional discrete-choice models usually assume that consumers are aware of all available alternatives in the marketplace and they make decisions out of identical choice sets. However, previous studies have shown that choices are often limited to a subset of all alternatives and that brand awareness varies across consumers (Draganska and Klapper 2011). For example, the survey data by Draganska and Klapper (2011) reveals that most ground coffee shoppers are only able to recall two to four brands out of the five major national brands. In our case, the online air-travel industry is composed of dozens of websites heterogeneous in market share and investment in advertising. Therefore, it will be even harder for the consumers to be aware of all possible choices and thus, making the assumption of full information set more problematic. Failure to account for the heterogeneity in consumer’s awareness sets will lead to biased estimates of brand preferences and responses to marketing activities (Goeree 2008). Therefore, more practically we allow awareness set to vary across consumers and model the formation of each individual at each purchase session explicitly in the first stage.

We use $A_{it} = (a_{i1t},...,a_{iJt}) = \{0,1\}^J$ to indicate consumer $i$’s awareness set at time $t$ for the total $J$ competing websites in the industry where $a_{ijt}$ denotes each site. We assume that awareness set membership is independent across websites, so the probability that consumer $i$’s awareness set equals $A_{it}$
is given by \( \Phi_{it} = \Pr(A_{it}) = \prod_{j \in A_{it}} \phi_{ijt} \prod_{j \notin A_{it}} (1 - \phi_{ijt}) \), where \( \phi_{ijt} \) is the probability that \( a_{it} = 1 \). We model a websites’ probability of belonging to the awareness set as

\[
\phi_{ijt} = \frac{\exp\left(\lambda_{0,ijt} + I_{ijt}\lambda_i\right)}{1 + \exp\left(\lambda_{0,ijt} + I_{ijt}\lambda_i\right)}
\]

where \( \lambda_{0,ijt} \) is a website-specific constant varying across individuals and \( I_{ijt} \in \{0,1\} \) indicates whether the consumer \( i \) has any past interaction (visit or purchase) with website \( j \) before the current purchase session. This specification allows a consumer’s awareness set to be affected by both population-level popularity of the website as well as individual-level history with the website.

All websites that are visited by the consumer later in the second stage should be included in her awareness set automatically because the consumer can only visit the websites that she is aware of. Therefore, these websites enter into the likelihood function as \( \phi_{ijt} \). However, it is not observable to researchers whether the consumer is aware of any other websites that are not visited. Therefore, we need to account for all possible combinations. If there are \( J \) competing websites and the consumer searches \( s \) of them, then there will be \( 2^{Js} \) potential awareness sets. If \( J \) is large is \( s \) is small, it will be infeasible to compute all the possible probabilities since the number of potential awareness sets increases exponentially. To address this issue, we adopt the method proposed by Goeree (2008) that simulate the choice set facing consumer \( i \) first and thereby reduce the relevant number of awareness sets to only one.

**Stage 2: Alternative Evaluation**

A consumer begins to visit websites for information search and alternative evaluation in the second stage after recognizing a purchase need. We use \( S_s \) to indicate the set of websites that are visited by the consumer in this stage. \( S_s \) is a subset of \( A_{it} \) because the consumer has to be aware of the website to visit the website.

We posit that consumer \( i \)'s visit to a specific website depends on the perceived utility, which is affected by her prior interactions with the website including ad clicks, visit and purchase history. The utility of visiting website \( j \) is given by:

\[
U_{ijt} = \bar{U}_{ijt} + e_{ijt} = \alpha_{0,ijt} + x_{ijt-1}\alpha_i + e_{ijt},
\]

Where \( x_{ijt-1} \) denotes all the factors that may impact consumer \( i \)'s perceived utility of visiting website \( j \) at time \( t \), \( \alpha_{0,ijt} \) captures consumer \( i \)'s intrinsic preference for website \( j \) that cannot be explained by \( x_{ijt-1} \), and \( e_{ijt} \) follows a generalized extreme value distribution. The coefficients \( \alpha \)'s are assumed to follow Multivariate Normal distribution to account for individual heterogeneity in website preferences.

The search theory asserts that a consumer starts his/her search process with the alternative with the highest utility (Ratchford 2008). Given the utility specification, the conditional probability to observe website \( k \) being chosen as the entry site by consumer \( i \) at time \( t \) given his/her awareness set \( A_{it} \) and individual parameters is given by

\[
Pr(F_{it} = k|A_{it}, \alpha_i) = \frac{\exp(\alpha_{o,ik} + x_{ik,t-1}\alpha_i)}{\sum_{k' \in A_{it}} \exp(\alpha_{o,ik'} + x_{ik',t-1}\alpha_i)}
\]
Stage 3: Purchase Stage

In the final stage of the purchase funnel, the customer decides which website to purchase from. Customers have to visit the website before they can purchase from it, therefore, the choice is made among visited websites in contracts to all websites in the market or the consumer’s awareness set. As we directly observe the entire search process made by each consumer, we are able to capture everyone’s choice set without any uncertainty. There is no need to model the heterogeneity in consumer’s choice set in a probabilistic manner like stage 1. In addition, as has been explained in the “Data” section, final purchase at one of the competing websites is required for a session to be included in our estimation sample. Therefore, no outside option is allowed in the purchase stage.

The utility of purchasing from website $j$ is given by the formula as follow:

$$U_{ijt} = \bar{U}_{ijt} + \epsilon_{ijt} = \gamma_{0,il} + x_{ilt}\gamma_i + \xi_{ijt}, \quad (5)$$

Where $x_{ilt}$ are the covariates that may affect consumer’s purchase utility and $\gamma_{0,il}$ is consumer $i$’s intrinsic website-specific preference that cannot be explained by other covariates. $\gamma_{0}$’s are assumed to be Multivariate Normal distributed to allow for consumer heterogeneity in preferences. $\xi_{ijt}$ is the unobserved part of utility and follows a generalized extreme value distribution.

Given the consumer $i$’s searched set at purchase session $t$: $S_l = (s_{i1t},…,s{iJt}) \in \{0,1\}^J$ and her individual tastes, the conditional probability of purchasing from website $l$ is given by

$$\Pr(P_{it} = l | S_{it}) = \frac{\exp(\gamma_{0,il} + x_{ilt}\gamma_i)}{\sum_{l' \in S_l} \exp(\gamma_{0,l'i} + x_{ilt}\gamma_i)} s_{ilt}. \quad (6)$$

In particular, if $s_{ilt} = 0$, the probability that website $l$ is chosen turns zero. If the consumer visits only one website, the probability that the entry site is chosen is one.

Covariates

The covariates that may affect the consumer’s utility of visiting and purchasing from a website are grouped into three blocks. The first block involves the number of visits to the website through different online channels, i.e., number of ad clicks. Whenever a consumer visits one of the competing websites through a redirecting website, we label the visit as one of the four following ad channels: 1) search engine, in which the consumer enters a keyword and redirects to the website by clicking either a sponsored or organic link (for which our data does not differentiate); 2) email, in which the consumer links to the website by clicking an email from his/her mailbox; 3) display ads, which we define as third party website that runs banner ads or video ads; 4) referral engines, which we define as third party website that provide a list of websites that competes in the same industry and may or may not provide reviews or ratings on the websites, such as TripAdvisor.com. If one visit does not have a redirecting website, we label it as direct visit, which means the consumer visits the website by typing in the URL directly in the browser. We then count the number of visits through different ways to every searched website in a purchase session.

We posit that the number of clicks on different ad channels in the previous purchase session affects consumer’s utility of visiting a website, while the number of ad click in the current purchase session affects the utility of purchasing on a website. This is because the more a consumer interacts with a website through an advertising channel, the more cumulative information stock she may accrue about the website. The information stock collected through different channels can take many functional forms. To simplify the estimation, we simply assume that it is a linear function of the number of visits to the website through a specific channel. But our model can be readily adapted to incorporate other formulations of the latent information stock.
The second block of covariates is consumer’s long-run history with the website, which involves the decayed cumulative amount of time and money spent on the website in the past. The decayed cumulative time (money) spent on the website $j$ at purchase session $n$ is calculated as

$$\text{Cumulative}_{\text{time(money)}}_{ijn} = \sum_{h=1}^{n-1} 0.7^{t_{ijn} - t_{ijh}} \text{time(money)}_{ijh},$$

where $t_{ijn} - t_{ijh}$ is the elapsed time (in month) between two purchase sessions and $\text{time(money)}_{ijh}$ is the time (money) spent on browsing the website at purchase session $h$. The decay rate 0.3 indicates that only 50% of information is retained after two months. We tried other decay rates and the estimation results do not change qualitatively. The consumer learns about a website by browsing and purchasing on it. The more s/he uses the website, the lower the costs of making sense of information sources and thing about the information gathered will be, which in turn lead to reduced time in finding relevant information and a faster check-out process (Shugan 1980). Therefore, cumulative time and money spent on the website will increase its attractiveness and thus, result in a higher utility of search and purchase.

The third block of covariates is consumer’s short-run history with the website – whether the consumer visited and purchased from the website in the last purchase session – or the so called state dependence. Consumer’s past search and purchase can lead to an intrinsic increase in preference for a website if structural state dependence exist, thus increase the search and conversion probability. We include two lagged indicators, one for search and one for purchase in the model to account for the state dependence of consumer choice. The distribution of preference parameters is fully accounted for, so if the state dependence coefficient is positive, it indicates that the website visit choices exhibit structural state dependence (Dubé et al. 2008).

There are two alternative explanations for inertia in search loyalty versus learning. Consumers may be unfamiliar with the website in terms of its layout (ease of use) or the potential benefit (discount, loyalty program, etc.), in which case, visiting the website provides relevant information. If learning, instead of loyalty, is the cause of state dependence, the effect of state dependence should diminish as the consumer accumulates more browsing experience with the website, because there is less and less information to be learned. The cumulative time spent on the website serves as a good proxy for website-level browsing experience. Therefore, we include an interaction term of the state dependence variable and cumulative browsing time as a means to test for the learning explanation. Under the learning assumption, the coefficient of the interaction term should be negative, which implies that state dependence is reduced as consumers’ brand experience increases.

For the purchase stage, we also control for the website’s position in the sequence of search process. In particular, we control for the first and last website visited during the current purchase session and use the in-between ones as baseline.

**Endogeneity and Instruments**

In our model, unobserved purchase utility and the number of ad clicks in the current purchase session may be correlated because consumers who are more likely to convert on the websites are those who are exposed to more advertisements as well, and may click the ads more times than others.

In order to address the potential endogeneity issue, we adopt the control function approach and use 1) the customer’s number of visits to the website through the same channel during the past 3 months except the visits due to the purchase of air tickets, and 2) the total number of visits to the website through the same channel by all panelists during the past 7 days as ad-click instruments for each channel. Instrument 1 reflects the consumer’s ad click propensity, thus is correlated with number of clicks on flight-related ads. But such propensity is with other product categories offered by the website, thus is not directly linked to the utility of the focal purchase. Instrument 2 reflects the website’s investment in a particular ad channel in the short-run and the higher it is, the more likely that a consumer will see an ad and click on it. However, the investment is not directly correlated with the attractiveness of purchasing on the website, because purchase decision is normally affected by price, product availability or customer loyalty instead of how heavy the advertising effort is.
We assume that the endogenous variables are additive in their observed and unobserved covariates. Therefore, the number of clicks made by consumer $i$ on website $j$ through channel $k$ at time $t$ can be written as a function of all exogenous variables entering the purchase utility as well as the two instruments discussed above:

$$ad_{ijkt} = W(x_{ijt}, z_{ijkt}; \psi) + \mu_{ijkt} \tag{8}$$

where $\mu_{ijkt}$ is the control function for channel $k$ and it affects $ad_{ijkt}$ but is independent of the unobserved purchase utility.

**Likelihood Function**

We assume that all the website-specific intercepts follow a Normal distribution, in order to capture unobserved heterogeneity in consumer preference while letting all other coefficients be identical for each individual. In addition, one’s awareness about a website and the likelihood to search can be correlated with some unobserved common shocks such as website design, offline advertising, etc. We tackle with this by allowing for correlations in random parts of the (website) intercept of the first two stages. We assume that $\lambda_{0,i}$ and $\beta_{0,i}$ are jointly distributed as

$$N\left(\begin{bmatrix} \lambda_{0,i} \\ \beta_{0,i} \end{bmatrix}; \begin{bmatrix} \sigma_{\lambda_{0,i}}^2 & \sigma_{\lambda_0 \beta_0} \\ \sigma_{\lambda_0 \beta_0} & \sigma_{\beta_{0,i}}^2 \end{bmatrix} \right).$$

Let $\vartheta = \{\sigma_{\lambda_0}, \sigma_{\beta_0}, \sigma_{\alpha_0}, \sigma_{\gamma_0}, \sigma_{\lambda \beta} \}$ denote all the standard errors to be estimated. The unconditional probability of observing consumer $i$'s purchase series is given by

$$\Pr\left(P_i = (l_{i1}, \ldots, l_{iT}) \right) = \prod_{t=1}^{T_i} \Pr\left(A_{it} \right) \Pr\left(F_{it} = k \mid A_{it} \right) \Pr\left(S_{it-1} \mid F_{it} = k, A_{it} \right) \Pr\left(P_{it} = l_{it} \mid S_{it} \right) f(\vartheta) d\vartheta \tag{9}$$

If $S_{it-1} \cap F_{it} = \emptyset$, then the probability of purchasing from the entry site is one and thus there is no need to estimate the following stages. Otherwise, the consumer chooses from all searched websites in stage 3.

The complete likelihood function to be maximized is

$$L = \prod_{i=1}^{N} \Pr\left(P_i = (l_{i1}, \ldots, l_{iT}) \right). \tag{10}$$

We use simulated maximum likelihood method to estimate all the parameters in our model next.

**Empirical Analysis**

**Parameter Estimates**

Table 9 presents the parameter estimates using the MSL estimator. The number of random draws taken is set to be 1000.

**Awareness Stage**

Our result shows that more than 99.9% of consumers in our sample have more than 50% chance to be aware of the following websites even without prior interaction with them: airline direct, Expedia.com, Priceline.com, Orbitz.com, and Travelocity.com. In contrast, more than 99.9% consumers have a high probability of not knowing about small agents. This result is consistent with these websites' number of searches and purchases made by the consumers in the sample. The coefficient for indicator of prior visit (10.128) is positive and significant, meaning that the consumer’s prior interaction with the website significantly increases the awareness probability in the current purchase session. This is even true with the case of small agents, because the magnitude of $I(\text{prior visit})$ is larger than that of small agents.
4.069). But we need to notice that even with prior visit, we still cannot assert that one website is sure to be included in the awareness set. This is because consumers have imperfect recall each time they perceive a purchase need and may forget some stores or brands they had past experience with.

**Alternative Evaluation Stage**

This stage is composed of two sub-decisions: 1) choose an entry site among all websites in the awareness set, 2) decide whether or not to visit other websites in the awareness set. The intercepts in the first sub-decision capture the relative unobserved attractiveness of all competing websites. We find that even though small agents have a low probability of being aware of, it enjoys a higher preference than any other competing websites once it is included in the awareness set. Consumers also favor airlines’ direct retail websites, expedia.com and Travelocity.com compared to other websites. As shown in Table 9, decayed cumulative browsing time increases a website’s probability of being chosen as the entry site. This could be due to the reduction in cost of re-visiting because consumers learn about the website by spending more time on it in the past and become proficient in using the website in the future. We also find that conversion on the website in the last purchase session does not affect the choice of entry site, but visiting the website last time increases the choice probability, which means consumers exhibit inertia in search behavior. Moreover, such inertia is due to an intention to learn more about the website because the interaction term between lag search and cumulative browsing time is negative and significant, which indicates that state dependence has a smaller impact on the choice of entry site as the consumer accumulates more browsing experience with the website and there is less information to be learned. Last but not least, we find that three out of four advertising channels have positive impact on the choice of entry site, in which search engine and display ads are most effective. However, their marginal effectiveness is not as strong as the number of visits to the website by typing in the address directly, because the later indicates the strongest familiarity with the website.

**Purchase Stage**

The first two columns Table 8 display the estimates for the purchase stage. We find that the cumulative amount of money spent on the website increase the conversion probability significantly, while the cumulative amount of time spent on browsing the website has no significant impact. We also find that if the consumer visits more than one website, being the first website visited does not increase the conversion probability, but being the last one does. Our findings also reveal interesting state dependence patterns. Once the customer reaches the third stage, the state dependence in search does not exist anymore as it does in the alternative evaluation stage. This means previous visits to the website only affect search decision, but has no significant impact on the purchase behavior for those deal seekers. However, state dependence in purchase still persists, in that last purchase increases the chance of current purchase.

The information stock gained by visiting the website through various channels also affects the conversion choice. Among the four advertising channels, display ads exert the largest impact on the choice of conversion site, followed by search engine. Visits through the other two ad channels do not appear to affect purchase probability significantly. We analyze the marginal impact of ad channels on conversion probability implied by the coefficients in the next session.

We also compare the estimates of the purchase stage of our proposed model with two benchmark models that do not consider state dependence or the entire funnel. The third and fourth columns of Panel C of Table 8 present the estimate of the benchmark model 1 in which inertia and history with the websites are not considered. We find that omitting these variables will not change the coefficients qualitatively, but it does lead to an over-estimation of ad effectiveness for search engine and display ads. The difference should be attributed to consumer’s intrinsic brand loyalty due to past interactions with the website or offline advertising. The fifth and sixth columns of Panel C of Table 8 display the estimates of benchmark model 2 that only considers the purchase stage. In this model, we mimic the scenario that which websites have been visited during the search stage are unknown, so the consumers make purchase decisions from all eight alternatives whereas in reality, they can only purchase from a website that has been searched. There are two main differences in findings according to this model. First, the effect of state dependence in search is significant, which means previous visit to the website affects current purchase decision. This model also shows that being the entry site significantly increases purchase probability. These two changes can be contributed to the ignorance of a large group of consumers who only search one website and
convert on the entry site. But this benchmark model confirms our findings on the effectiveness of advertising: only search engine and display ads have significant impact on conversion probability and display ads have a stronger effect. In general, our model fits the data much better than the two benchmark models as its log-likelihood is the lowest. Modeling the consideration set is more important than inertia because benchmark model 2 fits the data better than benchmark model 1. These comparative analyses demonstrate that incognizant of the intrinsic state dependence and the purchase funnel of consumers can not only lead to inferior models, but more seriously, managerial conclusions drawn from such popular benchmarking model can be erroneous and misleading.

Table 8: Parameter Estimates

<table>
<thead>
<tr>
<th>Awareness Stage</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>airline direct</td>
<td>7.454</td>
<td>0.032</td>
</tr>
<tr>
<td>cheaptickets.com</td>
<td>11.543</td>
<td>10.023</td>
</tr>
<tr>
<td>expedia.com</td>
<td>7.503</td>
<td>0.319</td>
</tr>
<tr>
<td>hotwire.com</td>
<td>4.387</td>
<td>14.719</td>
</tr>
<tr>
<td>orbitz.com</td>
<td>13.502</td>
<td>4.362</td>
</tr>
<tr>
<td>priceline.com</td>
<td>14.463</td>
<td>3.293</td>
</tr>
<tr>
<td>small agents</td>
<td>-4.069</td>
<td>0.764</td>
</tr>
<tr>
<td>travelocity.com</td>
<td>14.335</td>
<td>3.674</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alternative Evaluation Stage</th>
<th>Entry Site</th>
<th>Other Aware Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>airline direct</td>
<td></td>
<td>-10.639</td>
</tr>
<tr>
<td>cheaptickets.com</td>
<td>-0.507</td>
<td>-7.376</td>
</tr>
<tr>
<td>expedia.com</td>
<td>0.001</td>
<td>-4.020</td>
</tr>
<tr>
<td>hotwire.com</td>
<td>-0.195</td>
<td>-4.118</td>
</tr>
<tr>
<td>orbitz.com</td>
<td>-0.401</td>
<td>-3.172</td>
</tr>
<tr>
<td>priceline.com</td>
<td>-0.165</td>
<td>-2.970</td>
</tr>
<tr>
<td>small agents</td>
<td>2.226</td>
<td>-1.812</td>
</tr>
<tr>
<td>travelocity.com</td>
<td>-0.037</td>
<td>-3.166</td>
</tr>
<tr>
<td>Decayed cum. spending ($100)</td>
<td>-0.013</td>
<td>0.161</td>
</tr>
<tr>
<td>Decayed cum. browsing time (10 min)</td>
<td>0.079</td>
<td>-0.179</td>
</tr>
<tr>
<td>Lag purchase (0/1)</td>
<td>-0.049</td>
<td>-1.603</td>
</tr>
<tr>
<td>Lag search (0/1)</td>
<td>2.535</td>
<td>2.040</td>
</tr>
<tr>
<td>LagSearch*CumBrowsing</td>
<td>-0.081</td>
<td>-0.145</td>
</tr>
<tr>
<td>No. of visits: search engine at (t-1)</td>
<td>0.362</td>
<td>1.216</td>
</tr>
<tr>
<td>No. of visits: email at (t-1)</td>
<td>0.152</td>
<td>0.661</td>
</tr>
<tr>
<td>No. of visits: display at (t-1)</td>
<td>0.347</td>
<td>2.106</td>
</tr>
<tr>
<td>No. of visits: referral engine at (t-1)</td>
<td>0.062</td>
<td>0.957</td>
</tr>
<tr>
<td>No. of visits: self type-in at (t-1)</td>
<td>0.414</td>
<td>1.114</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Purchase Stage</th>
<th>Our Model</th>
<th>Benchmark 1</th>
<th>Benchmark 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>cheaptickets.com</td>
<td>-0.253</td>
<td>-0.370</td>
<td>-0.430</td>
</tr>
<tr>
<td>expedia.com</td>
<td>-0.379</td>
<td>-0.484</td>
<td>-0.437</td>
</tr>
</tbody>
</table>
**Ad Effectiveness**

Ads impact on conversion is two-folded: first, past ad-clicks increase a website’s probability of being chosen as the entry site or being searched, and thereby increase conversion probability indirectly; second, the number of visits through different advertising channels increases consumer’s purchase probability directly in the purchase stage. In the section, we compute the own- and cross- marginal impact of ad clicks on these two aspects to show the effectiveness of different advertising channels on different websites.

Modeling consumers’ choice among all competing websites allows us to analyze the self and competitive advertising effects for each website simultaneously. To compute the marginal impact of ad click on choice probability, we first calculate the simulated mean of individual-level intercept by

$$\bar{\beta}_i = \sum_r P\left(y_i | x_i, \beta^r \right) \beta^r, $$

where $\beta^r$ is a random draw from the population’s distribution and

$$P\left(y_i | x_i, \beta^r \right) = \prod_{t=1}^T \frac{e^{\beta^r x_{it} t}}{\sum_{j} e^{\beta^r x_{jt}}}. $$
Then we calculate the updated choice probability for each aware (or searched) website for every individual. Finally, we compute the average choice probability for each website across all consumers for ad click change with respect to each website.

Figure 2 Marginal Impacts of Ad Clicks on Probability of Being Chosen as Entry Site

Search engine

Email

Display Ads
Figure 2 depicts each website’s average probability of being chosen as the entry site given an additional visit to itself or other websites through search engine, email and display ads. The marginal impact of referral engine is not shown here because its parameter estimate is not significantly different from zero. The result indicates that one additional click on ad increases the website’s own probability of being chosen as the entry site, but the size of change varies a lot by websites and by ad channels. For example, an additional visit to small agents through search engine in the previous purchase session will increase its probability of being chosen as the entry site by 8.1%, but for cheaptickets, an additional click on search engine advertising only leads to an increase of 1.6%. Email ad-clicks have a smaller impact in general. For example, an additional click on email ad will only increase the probability for small agents by 3.3%. Our model structure also allows us to analyze the competitive pattern among all websites. From Figure 4, we also observe the competitive effects of search engine, email and display ads. For example, an additional visit to small agents decreases its competitors’ chance of being chosen by 1.1% on average. The effect size varies by website as well: airline companies’ retail website is affected the most with a decrease of 2.2%.

Figure 3 Marginal Impacts of Ad Clicks on Probability of Being Searched

Figure 3 shows the marginal impact of ad clicks on the probability of being searched for a website in the awareness set but not chosen as the entry site. Additional clicks on any advertisement increase search probability for all websites, but the magnitude of effects vary by ad channel and by website. For example, an additional visit through display ads to priceline.com in the previous purchase session increase its probability of being searched by 21.2%, but one more visit through email ads only increase the probability by 4.3%. In contrast, an additional visit through display ads only increase search probability for cheaptickets.com by 2.1%.
Search Engine

Display Ads

Figure 4 Marginal Impacts of Ad Clicks on Conversion Probability

Figure 4 shows the marginal impact of ad clicks on conversion probability. Only analysis for search engine and display ads are shown because the other two advertising channels do not have a significant impact on conversion probability according to our estimation results. We can do similar analysis here as what we have done for the choice of entry site. For example, an additional visit to cheaptickets.com through search engine increases its conversion probability by 5.14%, but one more visit to small online agents has barely any effect. We can also compute the cross marginal impact of ad click of one website on its competitors’ conversion probability. For example, an additional visit to airline companies’ direct websites through search engine decreases its competitors’ conversion probability by 1.4% on average. And the effect size varies by websites: expedia.com suffers the most with a decrease of 2.1%, while the conversion probability for small agents barely changes. In general, we find that websites with larger market share suffer more from ad-clicks on their competitors’ websites.

Conclusion

In this paper, we have developed an integrated three-stage model to measure the advertising effectiveness on consumer’s online search and purchase decisions in a multi-channel, multi-touch-point environment using individual-level advertising response data. We specifically model (1) consumers’ awareness, (2) consumers’ search by considering both their choice of entry site and their decisions about continuation of search, and (3) their subsequent purchases at one of the searched sites. Our model accounts for (1) the competitive effect by incorporating consumer’s choice among all relevant websites instead of a binary choice at one focal website, (2) state dependence effect, (3) heterogeneity across individuals and websites.

The results from our analysis shed several important insights on the effectiveness of multi-channel advertising. First, with our competitive dataset, we show that the marginal impact of ad clicks varies across websites. The impact of ad click on competitors’ conversion probability decreases as the competitor’s market share increases. Second, we show that different ad formats affect consumers differently based on the stage in purchase funnel. All four ad formats that we study have a positive impact on the search stage by increasing the probability of being chosen as the entry site or being searched. However, only search engine and display ads are effective in increasing the conversion probability directly. Third, we find that structural state dependence exist in both search and purchase stage. Ignoring state dependence will lead to severe over-estimation of ad effectiveness.
REFERENCES


Online Appendix: Simulation

We adopt the Maximum Simulated Likelihood method to estimate the three stages of our model jointly. MSL is consistent if the number of draws that are used in the simulation rises with sample size (Train 2009). Our model also allows for consumer heterogeneity in awareness set. Traditionally one has to compute all purchase probabilities that corresponds to each possible choice set, which leads to severe dimensionality problem as the number of alternatives increases. To solve this problem, we adopt the method proposed by Goeree (2009) that simulates the awareness set facing each consumer, thereby making only one purchase probability computation per individual necessary. The entire procedure of our simulation is introduced as follows:

1) A draw of $\mu_{ijr}$ is taken from a uniform distribution for each product-individual-time combination.

2) A draw of $\theta^r_i = (\lambda^r_{o,ij}, \alpha^r_{o,ij}, \beta^r_{o,ij}, \gamma^r_{o,ij}; \forall j)$ is taken from its assumed distribution for each individual. In specific, a set of random variables $\eta^r_j = (\eta^r_{\lambda, j}, \eta^r_{\alpha, j}, \eta^r_{\beta, j}; \forall j)$ are drawn from i.i.d $N(0,1)$. Then we can compute the website-specific intercepts in different stages for each individual as follows:

$$\lambda^r_{o,ij} = \lambda^r_{o,ij} + \sigma^2_{\lambda_{o,j}} \eta^r_j,$$

$$\alpha^r_{o,ij} = \alpha^r_{o,ij} + \sigma^2_{\alpha_{o,j}} \eta^r_j,$$

$$\beta^r_{o,ij} = \beta^r_{o,ij} + c_{j_2} \eta^r_{\beta_{2,j}} + c_{j_3} \eta^r_{\beta_{3,j}},$$

$$\gamma^r_{o,ij} = \gamma^r_{o,ij} + \sigma^2_{\gamma_{o,j}} \eta^r_j, \forall j$$

where

$$\begin{bmatrix} \sigma^2_{\lambda_{o,j}} & 0 \\ c_{j_2} & c_{j_3} \end{bmatrix} = \text{choi} \begin{bmatrix} \sigma^2_{\lambda_{o,j}} & \sigma^2_{\lambda_{\beta,j}} \\ \sigma^2_{\lambda_{\beta,j}} & \sigma^2_{\beta_{o,j}} \end{bmatrix}.$$

3) Calculate the probability that a website $j$ belongs to the awareness set $A_{it}^r$ for each individual-time combination:

$$\phi^r_{ij} = \frac{\exp\left(\lambda^r_{o,ij} + I_{ijt} \lambda^r_t\right)}{1 + \exp\left(\lambda^r_{o,ij} + I_{ijt} \lambda^r_t\right)}.$$

4) Given $\phi^r_{ij}$ and $\mu^r_{ij}$, construct a $J$ dimensional Bernoulli vector, $A_{it}^r$, which defines the awareness set for the $r$th loop. The $j$th element is determined according to

$$a^r_{ijt} = \begin{cases} 1, & \text{if } \phi^r_{ijt} > \mu^r_{ijt} \\ 0, & \text{if } \phi^r_{ijt} \leq \mu^r_{ijt} \end{cases}.$$

5) Calculate the probability of observing website $k$ as the entry site for each individual-time combination:

$$\Pr\left(F_{it} = k | A_{it}^r\right) = \frac{\exp\left(a^r_{o,ik} + x_{itk,t-1} \alpha^r_t\right)}{\sum_{k' \in A_{it}^r} \exp\left(a^r_{o,ik'} + x_{itk',t-1} \alpha^r_t\right)}.$$
6) Calculate the probability of observing searched set (excluding the entry website) for each individual-time combination:

\[ \Pr(S_{it} \setminus F_i | F_{it}, A_{it}^r) = \prod_{m \in S_{it} \setminus F_i} \frac{\exp(\beta_{0,il}^r + x_{im,t-1} \beta_{1}^r)}{1 + \exp(\beta_{0,im}^r + x_{im,t-1} \beta_{1}^r)} \cdot \prod_{m' \in A_{it} \setminus S_{it}} \frac{1}{\exp(\beta_{0,iim'}^r + x_{im',t-1} \beta_{1}^r)}. \]

7) If \( S_{it} \setminus F_i \neq \emptyset \), calculate \( \Pr(P_{it} = l | S_{it}) = \sum_{l' \in S_{it}} \exp(\gamma_{0,il}^r + x_{ijt} \gamma_{l'}^r) \); otherwise, \( \Pr(P_{it} = l | S_{it}) = 1. \)

8) Calculate

\[ L_{it}^r = \prod_{j \in A_{it}} \phi_{jjt} \prod_{j \notin A_{it}} (1 - \phi_{jjt}) \cdot \Pr(F_{it} = k | A_{it}^r) \cdot \Pr(S_{it} \setminus F_i | F_{it}, A_{it}^r) \cdot \Pr(P_{it} = l | S_{it}). \]

9) Repeat steps 1 to 8 many times and calculate the simulated probability as \( \hat{P}_{it} = \frac{1}{R} \sum_r L_{it}^r \), where \( R \) is the number of draws.

10) Calculate the simulated log-likelihood: \( SLL = \sum_{i=1}^{N} \sum_{t=1}^{T} \ln \hat{P}_{it} \).