Marketing Science Institute Working Paper Series 2012
Report No. 12-115

Attribution Modeling: Understanding the Influence of Channels in the Online Purchase Funnel

Hongshuang (Alice) Li and P. K. Kannan

Report Summary

Online marketers invest significant resources in driving traffic to their websites through multiple marketing interventions and channels. Authors Li and Kannan propose a conceptual framework to examine the nature of carryover and spillover effects of prior visits through different channels to a firm's website across a number of commonly used online channels, both at the visit and purchase stages.

They develop a three-level measurement model of customers' consideration of online channels/ sources, their visits through these channels over time, and subsequent purchase at the website, which accounts for carryover and spillover of visits. Based on customers' path data of visits and purchases at a travel and hospitality firm's website, they find significant carryover and spillover effects—for example, e-mails and display ads trigger visits through other channels, while e-mail leads to significant purchases through search channels. Attributing credit for the different channels for the purchase conversions using the model estimates, they find that the relative contributions of these channels are significantly different as compared to estimates of the widely-used "last-click" attribution model. A field study conducted at the firm's website by turning off paid search for a week validates the ability of the model in estimating the incremental effect of a channel on conversions.

This study highlights that the commonly used last-click attribution or the averaging attribution models are not good metrics for understanding the real impact of firm-initiated channels as well as customer-initiated channels on conversions.

Some of these metrics consider only those visits that result in conversion immediately to the exclusion of the other visits, and all of them consider only the visits that result in conversion while excluding the information that can be derived from non-conversions. While they may provide passable results in product categories with very short purchase funnels (with one or two touch points) and with fewer channels, they will invariably be misleading in product/service categories with longer purchase funnels as in high involvement categories—such as consumer durables and travel services—and for firms with multiple channels, both customer-initiated and firm-initiated. In the latter case, the effectiveness of firm-initiated efforts is generally underestimated using the last-click model.

In addition, there are limits to the effectiveness of firm-initiated efforts such as e-mail and display retargeting campaigns. Repeated e-mail targeting may actually hurt conversions in certain cases and retailers should be well-advised to carefully calibrate the targeting of customers with these instruments.

Hongshuang (Alice) Li is a doctoral candidate in the marketing program and P. K. Kannan is Ralph J. Tyser Professor of Marketing Science at the Robert H. Smith School of Business, University of Maryland.

Acknowledgment
The authors acknowledge and thank the support of Marketing Science Institute through their grant #4-1680 for this research.

Introduction

Firms use a variety of online marketing interventions such as advertising display and e-mail marketing to draw in potential customers to visit their websites. Customers also visit the website on their own initiative directly or through different sources such as search engines and referral sites. Upon customer's response (such as clicking on display ads, e-mail links, firm's paid search ads, or choosing any other source on their own), these interventions become the touch points or "channels" through which they visit and convert at the firm's website (Martin 2009; Mulpuru et al. 2011).

In practice, the effectiveness of campaigns across various touch points or channels is measured at the aggregate level using click-through-rates, which are, in turn, used to determine the level of investment (e.g., bids for search keywords) in each marketing campaign. However, such aggregate measures do not take into account the timing and sequence of earlier interventions, nor reflect their relative incremental impact in leading to website visits and conversions. Thus, what the aggregate measures suggest with regard to the effectiveness of these channels could be biased and misleading (Martin 2009). In addition, these channels are usually managed and measured using separate systems and often by different teams within an organization – display and paid search by one, e-mail campaigns by another, etc. (Green 2008), with incompatible data across different sample frames (Atlas 2008). The incompatible measurements across multiple channels results in double counting and disproportional attribution of conversion credit to each channel. Traditionally, a conversion at the website is credited to these different channels on the basis of "last-clicked" or "last-touched" channel, entirely ignoring the multiple channels a customer might have touched in the purchase funnel preceding the last click. However, for lack of a better measure, last-click conversion has become a standard attribution metric for determining advertisement payment and ROI of marketing interventions over the last decade (Latham 2011).

In many product and service categories, customers visit a firm's websites several times through multiple channels before a conversion occurs. A visit to the firm's website through a specific channel exposes customers to additional information regarding the attractiveness of the product and service vis-à-vis competing offers. The visit experience *per se* can have an impact on subsequent visits and conversions through the same marketing channel (i.e. carryover effects), or lead to visits and conversions through other channels (i.e. spillover effects). This impact can also be different across customers who tend to be very heterogeneous in how they respond to the various online marketing interventions (Mulpuru et al. 2011). Empirical study by Montgomery et al. (2004) finds that a customer's path of visits can reflect her goal, and the path information from as few as 6 previous visits can increase predicted purchase probabilities from 7% to 42%. Given this, the current practice based on "last-click" attribution can lead to biased and inaccurate estimates of ROI of marketing interventions which could then contribute to sub-optimal allocation of marketing budget across channels and campaigns. An integrated and accurate attribution model based on individual conversion paths is necessary to measure the contribution of multiple channels and overlapping campaigns and to assist decisions on optimizing marketing budgets. This is precisely the focus of our paper.

We develop a model for determining the nature of carryover and spillover effects of prior visits to a firm's website across a number of commonly used online touch points or channels, both at the stage of visiting the website and at the stage of purchasing at the website. Since not all customers may consider all channels in visiting a website (for example, some may consider search channels but are unaware of referral channels; some may be targeted by e-mails but others are not), accounting for the heterogeneity across consumers' consideration of channels is necessary to estimate the carryover and spillover effects without any biases. This leads to a three-level choice model that accounts for customers' consideration of channels through which to visit the website, the carryover, spillover and the sequence effects of prior channel interventions that contribute to the website visits, and the subsequent purchase conversions. The objective in building this model is to measure the incremental impact of a channel on conversions at a firm's website in a multi-channel context.

Our research is related to three streams of extant research with implications for carryover and spillover effects across channels. One examines customers' conversions *within* websites – focusing on the existence of lock-in effects within websites (Johnson, Bellman, and Lohse 2003; Zauberman 2003), learning effects impacting cognitive costs of using a website (Bucklin and Sismeiro 2003; Moe and Fader 2004) and the impact of demographic, site and visit characteristics (Danaher, Mullarkey, and Essegaier 2006). These studies do not account for the influence of a preceding channel visit or marketing interventions a visitor might have had before reaching the website that could affect the subsequent purchasing behavior, which is the focus of our study. The second stream of research focuses on the impact of individual channels *outside* the website such as display ads, e-mail and search engines in enabling conversions at the website (Chatterjee, Hoffman, and Novak 2003; Manchanda et al. 2006; Ghose and Yang 2009; Rutz and Bucklin 2011). Instead of focusing only on a specific marketing intervention as in the preceding work, our paper integrates the effects of a variety of marketing interventions/channels, such as search, display ads, e-mails, affiliate websites, referral engines, etc. on website visits and conversion (cf. Ansari, Mela, and Neslin 2008; Naik and Raman 2003). Finally, studies in the context of multi-channels have focused attention on customer lifetime value, total spending across channels and cross-selling, dynamics among earned media (Venkatesan and Kumar 2004; Kumar and Venkatesan 2005; Li, Sun, and Montgomery 2011; Kumar et al. 2012, Stephen and Galak 2012). However, none have examined the issue from the viewpoint of attributing credit for conversion to the multiple channels and understanding the impact of marketing interventions at different stages of the online conversion process. Our study fills this unique niche by being the first one to account for carryover and spillover effects across online channels, and to propose a methodology to apportion and allocate the credit for conversions that occur at firm's website to marketing channels.

We test our model using path data of visitors to the web site of an online firm in the hospitality industry. Our empirical analysis shows that there are significant carryover and spillover effects both at visit stage and purchase stage, the nature of which varies significantly across channels. For example, e-mails and display ads trigger visits through other channels, while e-mail leads to purchases through search channels. The empirical analysis also shows that our proposed model of attribution paints a much different scenario of relative contribution of these channels as compared to the widely-used last-click attribution model. For example, e-mail, display and referral channels contribute much more significantly to conversions than what the last-click model suggests while the contribution of search channels are significantly inflated compared to their real contribution. This has important implications for budgeting marketing investment across these channels. We also highlight the usefulness of our results through an illustration of whether or not the firm should retarget their customers with repeated e-mails using the visit path sequence. Finally, a field study conducted at the firm's website by turning off paid search for a week validates our model's ability to estimate the incremental effect of a channel on conversions.

In the next section, we present our conceptual framework. Section 3 presents the overview and details of the proposed model using a choice modeling framework. Section 4 provides details of the data and empirical results, the path sequence analysis for targeting, and the field study results. Section 5 highlights the managerial implications and contributions, and concludes with a discussion of limitations and future research.

Conceptual Framework

Our conceptual framework focuses on the purchase funnel in the context of online purchases of high involvement goods or services (see Figure 1, following Tables). The purchase funnel captures a series of stages that a customer moves through in making a purchase – the consideration stage, where the customer recognizes her needs and considers different channels for information search, the visit stage, where the customer visits the websites for information search and evaluation of alternatives, and finally the purchase stage, where the customer makes a purchase (e.g., Weisel, Pauwels, and Arts 2011). Given individuals' diverse habits for gathering information in the online shopping context, customers vary in their consideration of channels to use in visiting a firm's website. Some may be loyal to the website and consider going directly, while some may consider search channel for better prices and options. Some may consider both. We make a distinction between *customer-initiated channels*, which consumers seek out on their own initiative, and *firm-initiated channels*, where firms initiate marketing interventions such as display ads and e-mails (Bowman and Narayandas 2001; Weisel, Pauwels, and Arts 2011). The propensity to consider a customer-initiated channel might evolve over a long time horizon (Valentini, Montaguti, and Neslin 2011). Based on their awareness, experience, and expectations about these channels, they may make these channel consideration decisions in advance and store them in memory for use when the appropriate occasion arises. That is, consumers evaluate each channel they are aware of with regard to the benefit it provides versus the incurred search cost and arrive at a smaller set of channels they would consider for future information search when a purchase need arises (Hauser and Wernerfelt 1990; Mehta, Rajiv, and Srinivasan 2003). The channels in the consideration set act as "pre-decisional constraints" (Punj and Brookes 2002) to simplify the customer initiated search process when a purchase has to be made. On the other hand, the firm initiates marketing interventions targeting customers through e-mails and display ads. Customers conduct less pre-evaluation of these channels as compared to customer-initiated channels as the content of display ads and e-mails (such as specific promotions and deals) may vary from encounter to encounter. Thus, the firm initiated channel options enter into customers' consideration sets only when customers encounter them as a result of firm's targeting.

Conditional on their consideration sets, customers make visits to the firm's website through these channels and make a decision on purchase. Customers' prior visits have carryover effects in the same channel and spillover effects across channels both at the visit stage and purchase stage. We define carryover and spillover effects at the visit stage as the impact on the probability of a visit through a channel, while at the purchase stage we define them as the impact on the probability of making a purchase through a channel.

A customer's decision to visit the firm's website through a specific channel depends on the marginal benefits derived vis-à-vis marginal costs incurred in the visit. The benefit is the perceived mean attractiveness of making a purchase decision through the channel. The costs include the effort required to find the needed information (Shugan 1980) which can be viewed as an opportunity cost (Kim, Albuquerque, and Bronnenberg 2010) and the cognitive costs in processing the information (Johnson, Bellman, and Lohse 2003). As customers make multiple visits to the firm's website through various channels over time, the carryover and spillover of these visits increase or reduce these costs. As customers gain familiarity with a channel and its informational content, we expect the carryover of previous visits through that channel to reduce the costs in the same channel due to cognitive lock-in effects (Johnson, Bellman, and Lohse 2003; Bucklin and Sismeiro 2003), risk reduction over multiple visits, and self-reinforcement effects (Song and Zahedi 2005). The spillover across channels could reduce costs to the extent the channels are similar in nature and similar reinforcing information is sought by customers. If the channels are very different or if different types of information are sought by customers, then spillover could increase costs as customer may incur switching costs in breaking cognitive lock-in and adjusting to different types of channels. Thus, at the visit stage we model carryover and spillover through their impact on the costs of visiting a channel, with the costs reflecting not only the search cost, opportunity cost, and cognitive costs but also the mere exposure effects, reinforcement learning, and risk reduction as customer gather information across visits.

At the purchase stage, as customers make visits through different channels over time, the contextual information derived from the channels, such as information on other alternatives from a search engine or complementary goods from a referral site including their price and promotion, is compared and contrasted with the website's offering. This cumulative informational stock accrued over the past visits manifests itself as a utility of all prior visits through the channel, and gets added to the utility of the website's offering. Thus, the cumulative informational stock works to increase or decrease the overall utility of making a purchase at the website. The value of information gathered at a specific visit could decay over time depending on the channel and market dynamics, and thus the cumulative informational stock of prior visits would weigh the later visits more than the earlier ones (Ansari, Mela, and Neslin 2008; Terui, Ban, and Allenby 2011). Next, we develop our hypotheses.

Impact of carryover on the cost of visiting

For customer-initiated channels, as customers make repeat visits through a channel, the cognitive costs of visiting should go down due to cognitive lock-in effects (Johnson, Bellman, and Lohse 2003). So the cumulative experience (visits and time spent) in visiting through this type of channel should reduce the costs of visiting. On the other hand, since the firm-initiated channels – display ads and e-mails – may differ in content and specifics across different encounters, the impact of prior cumulative experience and exposure on a specific visit should be insignificant, unless the prior encounters were very recent and with the same content. However, Chatterjee, Hoffman, and Novak (2003) find that customers who are inclined to respond to display ad interventions would do so at their first exposure than later exposures. Thus, the carryover impact of firm-initiated contacts could hurt the future visits through the same channel. Also, DoubleClick (2004) reported a declining click-through on each additional display banner exposure. Ansari and Mela (2003) and Ansari, Mela, and Neslin (2008) suggest optimizing the content and timing of e-mail to maximize its impact by showing that e-mail, though costless to the firm, could generate negative influence on visits to the firm in the long run. Thus, we posit:

Hypothesis 1: *Carryover of visits through customer-initiated channels will reduce the costs of visiting through the same channel, while the carryover of impressions/ visits of firm-initiated channels will either have no impact or increase the costs of visiting in the same channels.*

Impact of spillover on the costs of visiting

Regardless of which channel a customer has experienced in the prior visit, when he/she encounters a targeted e-mail or display ad that provides very specific information on product/service features, price and promotion, it is likely to reduce the cost of clicking on the interventions and visiting the website. This is because content in the marketing interventions could be similar to the information that customer is seeking. Similarly, if a customer's prior visit to the website was through display ad or e-mail click-through, then the subsequent visit through any channel (especially a customer-initiated channel) is likely to be one where those specific product, price and promotion information are compared with other offers and information. Since the goal of such channel visits are clear with specific information requirement in mind, the spillover effect on the costs of visiting is also likely to be one towards reducing the costs of visiting. Also, Sherman and Deighton (2001) and Ilfeld and Winer (2002) report that banner exposure can increase ad awareness, brand awareness, and lead to more site visits. We expect that similar "billboard" effects could exist for e-mail interventions too. Information contained in e-mail newsletters can help customers refine their needs and narrow down their search domain. Also, firms can use e-mail campaign to direct customers to referral channels or direct channels that might be more lucrative to the firm (Myers, Pickersgill, and Metre 2004; Neslin and Shankar 2009). In addition, the product information and ongoing campaigns covered by search keywords and e-mail newsletters are very likely to overlap. Therefore, we can expect one channel to reduce the cost of visiting in the other.

With regard to spillover among search channels, referral channels and direct channels, when customers switch across these channels they are likely to search for complementary as well as comparative information on the product/service in the current channel vis-à-vis their prior channel. This could be more cognitive cost intensive as there could be switching costs due to different informational content and layout and the need to break cognitive lock-ins (Johnson, Bellman, and Lohse 2003). Customers will have to adapt to the new channel. As a result we can expect spillover to increase the costs of visiting. However, within search channels (organic versus paid search) we expect the spillover effect to reduce the costs of visiting as the informational content, layout, and experience effects to be reinforcing.

Hypothesis 2a: *The spillover of prior customer-initiated channel visits on the costs of visiting in marketing-initiated channels is negative (reducing the costs of visiting) and the* *spillover of prior marketing-initiated channel visits on the costs of visiting in customer-initiated channels is also negative.*

Hypothesis 2b: *While spillovers among organic and paid search channels will reduce the costs of visiting through the other, the spillover across search, referral, and direct channels are likely to increase the costs of visiting through the other two channels.*

Impact of carryover on purchase probabilities

Extant research suggests that display banner ad exposures seem to be processed at a pre-attentive level and may benefit ultimate purchase (Dreze and Hussherr 2003; Manchanda et al. 2006). Manchanda et al. (2006), using a hazard modeling approach find display ads can accelerate the purchase timing. In addition they find the number of display impression as well as the number of sites and pages containing the display ads all have a positive impact on the repeat purchase probability. A recent ComScore report also finds the banner ad impression could be more influential in leading to conversions than the click-throughs (Lipsman 2012). Thus, we would expect the carryover impact of display ads to be positive on purchases. A similar argument can be made with regard to e-mails, too. Repeated direct visits, which are customer initiated, could imply that a customer has a higher preference for the firm’s offering (Bowman and Narayandas 2001) and thus does not shop around in other channels. This carryover could lead to a positive impact on purchase probabilities. With regard to carryover of search and referral channels, one can expect that the customers visiting through these channels could be more price sensitive and focused on finding better deals. Yet, if a customer has made repeat visits to the websites through search and referral channels, it might indicate he/she finds the website’s offering to be more attractive as compared to the other ones they encountered in prior visits in search or referral channel, and hence is more likely to make a purchase (positive carryover). Chan, Wu, and Xie (2011), for example, show that the customers acquired through paid Google search channel make more purchases and generate higher customer lifetime value than customers acquired from other channels. Weisel, Pauwels, and Arts (2011) also find paid search has high profit impact. Compared with e-mail, the profit impact of paid search is more enduring, i.e. it wears in faster and wears out more slowly. Overall, the expectation for positive carryover is strong.

Hypothesis 3: *The carryover effects are positive on purchase probabilities.*

Impact of spillover on purchase probabilities

Yang and Ghose (2010) examine the spillover between organic search and paid search and report a positive asymmetric pattern, i.e. the impact of organic search on paid search is over three-times stronger than the impact of paid search on organic search. They also conducted field experiment to show that the total click-through rates, conversion rates, and revenue are lower in the absence of paid search than in the presence of it, highlighting the spillover from paid search. We could therefore expect positive spillover effects across search channels. With regard to firm-initiated channels, we should expect carryover of e-mail and display ads to have positive impact on purchase probabilities in any of the customer-initiated channels. Such repeat response to firm-initiated channels indicates higher preference level for the firm's offering leads to positive spillover and increase in overall purchase probabilities regardless of which channel they make a visit through (cf., Manchanda et al. 2006).

Hypothesis 4: *The spillover effects among organic and paid search channels on purchase probabilities are positive, and the spillover effects of firm-initiated channels on purchases through customer-initiated channels are also positive.*

We do not have a priori expectations with regard to the spillover effects across search, direct and referral channels, or the spillover from these customer-initiated channels on purchases through firm-initiated channels. We expect these effects to depend on the preference intensities and price sensitivities of customers visiting through these channels, and let the data provide us context dependent insights.

In sum, the above framework and hypotheses provide the basis for a measurement model where we use (1) costs as a catch-all to account for all the factors that impact a visit to the firm's website through different channels (both impediments and facilitators) and (2) cumulative informational stock to characterize the value of information gathered in prior visits through different channels relative to the firm's offering, which together impact the purchase probability during a visit.

Model Overview

As shown in Figure 1, the conversion decision of a customer at an online site consists of three stages: the consideration of alternative channels to use and/or the marketing interventions if encountered, the visit decision and the purchase decision. We develop an individual-level probabilistic model explicitly accounting for these stages.

Consideration of channels

Given the diverse individual habits in gathering information in the online shopping context, we expect to see a significant variation in customers' consideration of channels to use in visiting a firm's website. In order to control for individual heterogeneity in the consideration of channels, we allow individuals in our model to have different consideration sets of channels, which could include both customer-initiated channels and firm-initiated channels. We assume that an individual's consideration of channels to use in visiting the firm's website is the same across all visits and purchase occasions, except for firm-initiated channels (display ad and e-mail) which only enter into consideration when a customer has encountered them. This is because we deal with data collected in a short time window during which the firm's marketing strategies and tactics remained constant. Also, recent research findings in the context of web browsing and purchasing, support that consumers have fixed consideration sets, with size and elements being heterogeneous across customers (De Los Santos, Hortaçsu, and Wildenbeest 2012).

Assume there are Q channels available for a customer to reach the firm's website on their own initiative, and meanwhile, the firm operates *(J-Q)* firm-initiated channels. Thus, a customer's consideration set could include up to J channels.

To study the consideration of customer-initiated channels, we assume, following the model proposed by van Nierop et al. (2010), that individual i (i=1,…,I) has a Q-dimension vector of latent utility, $\tilde{C}_i^*$, for considering each customer-initiated channel q (q=1,…,Q) in the visit decision. The Q-dimension vector $\tilde{C}_i^*$ is jointly drawn from a multivariate Normal distribution as in Equation (1). Further, each element of latent utility c_{iq}^* is determined by a set of customer-specific characteristics R_i shown in Equation (2). The latent utility c_{iq}^* is associated with a binary value c_{iq}, where $P(c_{iq}=1)=P(c_{iq}^*>0)$ implies that channel q is included in individual i's consideration set. We normalize all the diagonal elements in Σ to be 1 for identification, so that the off-diagonal elements are, therefore, the correlations of considering two channels.

$$\tilde{C}_i^* = (c_{i1}^* \ldots c_{iq}^* \ldots c_{iQ}^*)^T \sim N_Q(\varphi, \Sigma) \qquad q = 1, \ldots, Q \tag{1}$$

$$c_{iq}^* = R_i \alpha_{iq} + \varepsilon_{iq} \tag{2}$$

For the firm-initiated marketing interventions, we use $\{c_{i(Q+1)}, \ldots, c_{iJ}\}$ to indicate whether customer *i* encounters any marketing intervention in channel *(Q+1)* to channel *J* in each of their visit decision.

We exclude the empty consideration set from our model (van Nierop et al. 2010), since we can observe a customer only if she has made at least one visit to the focal firm's website through one of the *J* channels. Define H_k as one combination of any positive number of channels out of J channels, where $k = 1, \ldots, (2^J - 1)$. The multivariate probit variable $C_i = (c_{i1} \ldots c_{iJ})^T$ is the same as H_k with a probability $P(C_i = H_k \mid \alpha, \Sigma)$.

Given the consideration of channels, we model the visit decision and subsequent purchase decision in a two-level nested logit framework. That is, the realization of the consideration set determines the structure of the nested logit model. At any online visit occasion *n* (n=1,...,N_i) , individual *i* can choose to visit the firm's website through channel *j*, (V_{in} =*j*, $j \in \{c_{ij} = 1\}$), searching for and gathering new information to possibly make a purchase, or not make any visit at all (V_{in} =0). Notice that channel *j* can be either a customer initiated channel considered for use ($j \in \{c_{ij} = 1,\ 1 \le j \le Q\}$) or a marketing intervention encountered ($j \in \{c_{ij} = 1,\ (Q+1) \le j \le J\}$). Given the visit through channel *j*, individual *i* may decide to make the purchase in the same visit, B_{ijn} = 1, or not, B_{ijn} = 0. We assume that some search precedes the purchase stage in every occasion *n*, because the consumer has to at least search for the availability of a specific service (e.g., airline seat availability on a specific date) before purchasing. Given the specific set of considered channels, C_i , the probability of purchase by individual *i* via channel *j* at occasion *n* is:

$$P(B_{ijn} = 1, V_{in} = j \mid C_i) = \Pr(B_{ijn} = 1 \mid C_i, V_{in} = j) P(V_{in} = j \mid C_i) \tag{3}$$

In the following, we first introduce the purchase decision and then discuss the visit decision, where the option value of a purchase is accounted for through inclusive values.

Purchase decision

Conditional on the consideration of and the visit through a certain channel, consumer *i*'s perceived utility of purchase in channel *j* at occasion *n* is W_{ijn} (Equation 4). The conditional purchase probability is determined based on a logit form (Equation 5), where τ is the scale parameter for the visit decision associated with the purchase decision. The error term ζ_{ijn} follows a generalized extreme value distribution. The utility of no purchase is $W_{i0t} = 0$.

$$W_{ijn} = \bar{W}_{ijn} + \zeta_{ijn}, \qquad j = 1,\ldots,\mathrm{J}, \tag{4}$$

$$\Pr(B_{ijn} = 1 \mid C_i, V_{in} = j) = \frac{\exp(\bar{W}_{ijn} / \tau)}{1 + \exp(\bar{W}_{ijn} / \tau)}, \qquad j = 1,\ldots,\mathrm{J} \tag{5}$$

In Equation (4), $\bar{W}_{ijn}$ is the overall attractiveness of making a purchase through channel *j*. We assume that overall perceived attractiveness of purchasing a product/service can vary along some mean attribute level of the offering (Erdem and Keane 1996). In our context, since the travel and hospitality service in every purchase is unique and distinct, and thus could be a new experience to the consumer, we construct a model where consumers are imperfectly informed about these attribute levels of the service. At the outset, consumer i perceives the mean attribute level of her target service to be purchased in channel *j* as γ_{ij} in Equation (6).

$$\bar{W}_{ijn} = \gamma_{ij} + \sum_{k=1}^{J} \gamma_{ij,k} G_{ikn}. \tag{6}$$

γ_{ij} is set by prior experiences and expectations of the attractiveness of purchasing through a channel. For example, a customer going to the firm's website through a click on display ad or an e-mail or through a coupon/referral site may have some mean expectation of the attractiveness of the purchases she might make. The overall attractiveness of making a purchase is then updated by the information she collects through channel visits, e.g., search engines (Google, Yahoo, etc.), referral engines (TripAdvisor.com, etc.) or the focal company's website and by the information conveyed in marketing interventions such as display ads and e-mails the customer may encounter. For each of the *J* channels, including *Q* customer-initiated channels (such as search, direct, and referral), and *(J-Q)* channels of firm-initiated marketing interventions (display ad, e-mail), the expected perceived overall attractiveness at occasion *n* is in Equation (6). The term G_{ikn} detailed in Equation (7) is the cumulative informational stock/content that contains the informational influence of all the preceding visits that individual *i* has been exposed to in channel *k* up to the $(n-1)^{th}$ visit, where $n=1,\ldots,N_i$ (Ansari, Mela, and Neslin 2008, Terui, Ban, and Allenby 2011). The indicator d_{ikh} equals to 1, if individual *i* chooses to visit channel *k* at occasion *h*. The informational effect of past channel visits decays at a channel-specific decay rate λ_k , according to the elapsed days $(t_{ikn}-t_{ikh})$. The instantaneous informational influence of any visit/intervention is normalized to 1, but the relative magnitude of this instantaneous influence of channel *k* compared with other channels can be picked up by the coefficients $\gamma_{ij,k}$ in Equation (6).

$$G_{ikn}=\sum_{h=1}^{n-1} d_{ikh}\times(1-\lambda_k)^{(t_{ikn}-t_{ikh})} \tag{7}$$

A visit to the website through a channel will incur cost, S_{ijn} ,which will be captured only in the visit decision, but treated as sunk cost in the purchase decision discussed in this subsection. In sum, consumer i 's expected utility of making a purchase in channel *j* at occasion *n*, W_{ijn} , is captured by Equation (8).

$$W_{ijn}=\bar{W}_{ijn}+\zeta_{ijn}=\gamma_{ij}+\sum_{k=1}^{J}\gamma_{ij,k}G_{ikn}+\zeta_{ijn} \qquad j=1,\ldots,\mathrm{J} \tag{8}$$

Visit decision

We posit that consumer *i*'s decision to visit channel *j* at visit occasion *n* depends on the perceived utility for that visit. The perceived utility U_{ijn} (Equation 9) is a function of customer *i*'s perceived benefits of visiting channel *j*, $\beta_{0,ij}$ (say, the useful information they can gather from the visit), and the attractiveness of the purchase/no purchase option through that channel on occasion *n* captured by the inclusive value term and its coefficient, τI_{ijn}, minus the disutility of the incurred cost $\beta_{ij}S_{ijn}$. Consumer *i*'s inclusive value of purchase or no purchase option in channel *j* at occasion *n* is $I_{ijn}=\log\left\{1+\exp(\bar{W}_{ijn}/\tau)\right\}$. The error term η_{ijn} follows a generalized extreme value distribution. The utility of not visiting, U_{i0n}, is normalized to be 0. At each visit occasion, the customer compares the perceived net utility of visiting by trading-off the potential purchase benefits against the incurred costs, and chooses to visit the channel that offers the greatest net utility or not visit at all.

$$U_{ijn} = \bar{U}_{ijn} + \eta_{ijn} = \beta_{0,ij} + \tau I_{ijn} - \beta_{ij} S_{ijn} + \eta_{ijn} \qquad j = 1, \ldots, J \tag{9}$$

The cost S_{ijn} is further parameterized in a logit form as

$$S_{ijn} = \frac{\exp(\mu_j T_{ijn} + \sum_{k=0}^{J} \mu_{j,k} L_{ik,n-1})}{1 + \exp(\mu_j T_{ijn} + \sum_{k=0}^{J} \mu_{j,k} L_{ik,n-1})}, \tag{10}$$

Thus, cost is bounded in the range of [0, 1]. That is, it is always costly to make a visit, but total costs level off as the customer experience and knowledge in a channel reaches a certain amount. This functional form for total cost is also consistent with the empirical finding of Moorthy, Ratchford, and Talukdar (1997) that unit search cost is quadratic as a function of experience, with an initial increase and then a decrease. T_{ijn} is the cumulative time spent at website visiting through channel *j* capturing carryover of these visits. This is determined on the basis of the difference between the start time stamp and the end time stamp associated with each visit/impression. We also include a set of *(J+1)* lag visit dummies, $\{L_{ik,n-1}, k = 0, \ldots, J\}$, indicating the channel visited by consumer *i* at occasion *(n-1)*, with 0 representing no visit in last occasion. This can be viewed as a first order Markov process to capture the short-term carryover and spillover effects[1].

The coefficients in the cost function, μ_j and $\mu_{j,k}$'s, can be either positive or negative. For example, positive μ_j or $\mu_{j,k}$'s imply the corresponding variables can increase the cost S_{ijn}, while negative μ_j or $\mu_{j,k}$'s imply reducing the cost. In addition, the coefficients μ_j and $\mu_{j,k}$'s capture the relative importance of total previous visits in the same channel (long-term carryover) versus the latest visit through channel *k* (short-term carryover or spillover) to the total cost S_{ijn}. Meanwhile, the coefficient of cost, β_{ij}, in Equation (9) determines the relative disutility of the cost S_{ijn} compared to $\beta_{0,ij}$ and τI_{ijn} in the utility function. Thus, with this formulation, we can

[1] We use visits lagged by one period, based on previous findings by Montgomery et al. (2004) that the first order Markov performs better than zero order Markov process. This could also be viewed as behavioral reinforcement. In addition, our empirical application when we accounted for the visits in (n-2) occasion, it did not significantly change the relative magnitude of costs across channels. Neither did it improve the goodness of fit of the model.

compare the marginal impact of the cost of visiting with β_{ij} and compare the relative importance of long-term carryover versus last visit with μ_j and $\mu_{j,k}$'s. In order to identify the coefficient β_{ij} as well as μ_j and $\mu_{j,k}$'s, we set $\mu_{j,0}$'s to be 1. Thus, the visit decision is a comprehensive decision, because it takes into account not only the short-term impact of lagged visit $L_{ik,n-1}$ in S_{ijn} but also the long-term accumulated informational stock of past visits and marketing interventions in T_{ijn}, as well as the inclusive value terms, I_{ijn}.

Notice that consumer *i*'s visit decision is conditional on her consideration set. That is, given C_i, the probability of individual i visiting channel j at occasion n is

$$\Pr(V_{in} = j \mid C_i) = \frac{c_{ij} \exp(\bar{U}_{ijn})}{1 + \sum_{k=1}^{J} c_{ik} \exp(\bar{U}_{ikn})}, \qquad \text{for} \qquad j = 1, \ldots, J \tag{11}$$

and $$\Pr(V_{in} = 0 \mid C_i) = \frac{1}{1 + \sum_{k=1}^{J} c_{ik} \exp(\bar{U}_{ikn})}$$

Overall, taking into account the consideration, visit and purchase stages, the joint likelihood function in Equation (12) contains a set of parameters $\theta = (\alpha, \Sigma, \beta, \mu, \tau, \gamma, \lambda)$. The model is estimated using MCMC approach which provides a computationally tractable estimation of the large number of parameters in the model.

$$L(B \mid \theta) = \prod_{n=1}^{N_i} \prod_{i=1}^{I} \prod_{j=i}^{J} \sum_{k=1}^{2^J - 1} P(C_i = H_k \mid \alpha, \Sigma) \times \left[b_{ijn1}{}^{B_{ijn}} b_{ijn0}{}^{(1-B_{ijn})} \right] \tag{12}$$

where
$$\begin{aligned} b_{ijn1} &= P(V_{in} = j \mid C_i; \beta, \mu, \tau) P(B_{ijn} = 1 \mid C_i, V_{in} = j; \gamma, \lambda) \\ b_{ijn0} &= P(V_{in} = j \mid C_i; \beta, \mu, \tau)[1 - P(B_{ijn} = 1 \mid C_i, V_{in} = j; \gamma, \lambda)] \end{aligned}$$

For details of priors and full conditional distributions, please contact the authors.

Empirical Analysis

Data

The data for this study is provided by a travel and hospitality franchise firm in the travel and hospitality industry. The firm uses a variety of online marketing channels, such as e-mails, search engines – both organic and paid search – display ads, referral engines and affiliates, etc. to attract customer visits to their website[2]. The average monthly visit to the firm's website in 2010 was around 26 million. The path data for each customer is developed by integrating data feeds from DoubleClick for advertisers (display ad and search engines), Omniture Site Catalyst (visits from different sources using cookies and login IDs), affiliate websites, and e-mail campaign management system. Overall, the path data provides information on display impressions and e-mail drops each customer encountered over time and whether it was clicked or not, click through visits from search engine (organic and paid), referral sites, and direct visits[3]. It does not include visits to search engines and referral sites that did not result in a click-through to the firm's website but this is captured by the outside option in our model as they do not materialize in visits to the firm's website. The firm can also use cookies and login IDs to identify their rewards program customers and their specific rewards program tiers – Rewards Level-1, Rewards Level-2, Rewards Level-3 and Rewards Level-4, from the lowest level to the highest. Across tiers, there are differences in customers' purchase frequency as well as purchase funnel (Rewards Level-4 is given to individuals as honorary membership, not based on actual purchases).

The dataset is a random sample from visitors to the firm's website during a week in late August, in the past year, with their visit history between late June and late August. We track each visitor's 68 days' history containing whether an online visit was made each day, visits through different channels to the focal firm's website, the instances of marketing interventions, and purchases if any. In our data, the average time between the first visit since the last purchase and the current purchase was 9.2 days, indicating that a 2-month window should be sufficient to capture all relevant historical data to explain visit and conversion decisions. Among the 1997 customers sampled from the cohorts, 163 made multiple purchases ranging from 2 to 11 times). We applied stratified sampling to assure the overall and channel-wise conversion rates in the sample are close to the firm's average of 4.5% and to allow us reliably estimate the impact of various independent variables on conversion at the website. All contiguous visits through the

[2] Organic Search and Paid Search represent the visits originated from a click at search engines, such as Google, Bing and Yahoo. Organic search is free traffic to the firm's website, while paid search involves a fee per click for the firm. Referral engines include referral sites such as TripAdvisor.com and Kayak.com, B2B referrals, event management tools, social media~~.~~ E-Mail channel represents the visits by a visitor who received an e-mail and clicked the link embedded in the e-mail. It also includes visits from a guest who received an e-confirmation of their booking or pre-arrival e-mail and clicked through the link in that e-mail. Finally, Display channel represents those visits made to the website by clicking on a display banner advertisement.

[3] A Direct visit is made by customers via typing in the URL of the firm's website.

same channel within 30 minutes with the same campaign code are treated as a single visit. The summary statistics in Table 1 (Tables follow Reference throughout) are based on 1997 unique customers' data, comprising 22369 click-through visits to the firm's website. The Display channel in Table 2 includes the display impressions with no click-through by customers who have visited the firm's website. Overall, 815 customers made 1128 purchases over the study duration. As seen in Table 1, the conversion rates in each channel vary significantly with Display being the lowest and Paid Search being the highest.

Model fit

The proposed model is compared with alternative models on the dimensions of model fit and model predictions and outperforms all of them. Table 2 (Tables follow Reference throughout) provides the model fit details of the proposed model and alternative models in terms of Log Marginal Likelihood values and the mean absolute percentage error (MAPE) of fit using the calibration sample. The alternative models estimated include (1) Model 1, that has all stages but does not include the decay parameters in the information stock variables in the purchase stage (that is, decay is assumed to be zero for all visits), (2) Model 2, which contains only the visit and purchase stages (each consumer considers all channels – exogenously specified with no variations across customers), (3) Model 3, which has all of consideration stage, visits and purchases but does not include the lagged visits as explanatory variables in the visit stage, and (4) Model 4, a naïve model with only channel specific constants at the visit stage and purchase stage, and (5) the proposed model. The model fit in terms of the Log Marginal Likelihood values indicates that the proposed model is superior to the alternative models. Additionally, the results indicate that the consideration sets, the lag variables in the visit stage and the decay parameters in the purchase stage do play a significant role in contributing to the explanatory power of the model, and thus are important variables to consider in explaining visits and purchases at the firm's website. It is particularly noteworthy that the lag variables as part of costs in the visit stage contribute significantly to the fit of the model. Overall, the model comparison shows that our proposed model and the model lack of consideration stage (Model 2) perform reasonably well, while other models perform significantly inferior.

Model estimates

Table 3 (Tables follow Reference throughout) provides the estimates of the proposed model. These estimates are posterior means based on 5,000 MCMC iterations, after 20,000 iterations used as burn-in. We investigate the iteration plots and use Geweke convergence test (Geweke 1992) where we compare the estimated parameters based on the first 1000 iterations, the 2001-3000 iterations, and 4001-5000 iterations after burn-in period to determine the convergence to stationary posterior distributions of the parameters in the proposed model. The table shows the channel specific estimates for the four customer-initiated channels – Organic Search, Paid Search, Referral and Direct – and two marketing intervention based channels – E-mail and Display – at the consideration, visit, and purchase stages. We discuss these stages separately.

Consideration stage. We model a consumer's consideration of customer-initiated channels (Organic Search, Paid Search, Referral and Direct) as a function of their level of membership in the firm's loyalty program (non-member, Rewards Level-1 through Rewards Level-4). We expect the membership levels to act as a proxy for consumers' experience, affect and commitment towards the firm's brand and capture their impact on the channels they would consider in visiting the website. As shown in Table 3, a non-rewards-program member is more likely to consider Organic Search and Paid Search as compared to the rewards program members at any level, while they are less likely to consider Referral and Direct channels as compared to the rewards program members. Rewards Level-3 and Rewards Level-4 members are more likely to consider Direct as compared to the Rewards Level-1 and Rewards Level-2 members. The estimated correlation matrix of consideration (not reported) indicates that customers are more likely to consider Organic and Paid Search together (correlation coefficient .69) and Referral and Direct together (correlation coefficient .87). An analysis of the posterior distribution of the consideration set probabilities (not reported) indicates that non–rewards members (over 85% of them) consider all customer-initiated channels, while around 20% of Rewards Level-3 and Rewards Level-4 members consider only Direct channel with a small percentage of them (< 10%) considering all customer-initiated channels. Overall, we find a significant heterogeneity in the consideration of the customer-initiated channels.

Visit stage. The estimates of visit stage in Table 3 provide (1) the long-term carryover effects of prior visits on costs of visiting the channel through the inclusion of cumulative time spent visiting through each channel and (2) short-term carryover and spillover effects through the use of lag variables. The coefficients for cumulative time indicate that for all customer-initiated channels, except Organic Search, the carryover effects on the costs of visiting the channel is significantly negative (thus reducing the costs). This could be due to cognitive lock-in effects (Johnson, Bellman, and Lohse 2003), mere exposure effects, reinforcement leaning effects and risk reduction kicking in with increased experience in visiting through customer-initiated channels, thereby reducing costs of re-visiting as expected in Hypothesis 1. The long-term carryover terms of firm-initiated channels, however, are not significant. This is consistent with Chatterjee, Hoffman, and Novak (2003) and Doubleclick's (2004) results that customers who respond to display ad interventions would do so at their first exposure rather than later and repeated display ad exposures have no added impact.

The short-term carryover effects (Lag-Organic on Organic Search, Lag-Paid on Paid Search, and so on, ranging from -1.26 to -2.43) indicate that all these effects contribute to reducing the costs of re-visiting. That is, if a customer made a visit through a specific channel in the last occasion (within the last day or on the same day), the cost for the current visit through the same channel is reduced. The lag effects of Organic Search on both E-Mail (-.30) and Display channel (-.25), and the lag effects of Paid Search (-.49 on E-Mail and -.43 on Display) indicate a spillover effect of these customer-initiated channels in reducing costs of visiting through firm-initiated channels, i.e. E-Mail and Display, which is consistent with Hypothesis 2. However, spillover effects of firm-initiated channels on customer-initiated channels are, by and large, mixed. For example, prior Display visits reduce the costs of visiting through Organic and Paid Search, consistent with the findings of Sherman and Deighton (2001) and Ilfed and Winer (2002) which show that display ad exposure increase ad awareness, brand awareness, and lead to more site visits ("billboard effects"). On the other hand, the lag effect of E-mail visit increases the cost of visiting through Organic Search (.74), Direct visit (.24) and Display (.49). A possible explanation for this could be that those customers visiting the firm's website clicking through e-mails are more likely to come back through E-mail channel or shop around using Paid Search or Referral channels. As for the lag effect of Organic Search on Paid Search and vice-versa, the spillover effects reduce costs of visiting through the other channel. However, we find that the spillover effects of Paid Search on Organic Search (-.79) are much stronger than in the reverse direction (-.18). This is contrary to what Yang and Ghose (2010) find in their study that Organic Search has a much stronger effect in leading to clicks in Paid Search than the reverse effect.

The coefficients for the costs of visit vary across channels reflecting the extent to which the visit decisions in these channels are sensitive to these costs. The coefficients for Referral, Direct and E-Mail (-3.58, -3.11, and -3.58) are the highest in magnitude indicating that a unit drop in costs of visiting is likely to impact repeat visits through each of these channels much more significantly than that for the Organic Search, Paid Search and the Display channels. These results highlight that the impact of carryover or spillover could be much higher for Referral, Direct and E-Mail channels as compared to the other channels. Finally, the coefficient of the inclusive value is significant (.35, which is closer to 0 than 1) indicating that inclusive value plays a critical role in trading off the perceived attractiveness of the purchase/no-purchase option in a channel versus the incurred costs of visiting through that channel.

Purchase stage. At the purchase stage, the informational stock captures the impact of prior visits with their respective decays over time, indicating the lingering effect of information gathered in prior visits on purchase probability of the current visit. We find that the carryover effects of firm-initiated channels (E-Mail and Display) are significantly contributing to increase purchase probabilities. These results are consistent with extant research which suggests that the exposures to display banner ads seem to be processed at a pre-attentive level and may benefit ultimate purchase (e.g., Dreze and Hussherr 2003; Manchanda et al. 2006). Specifically, Manchanda et al. (2006) find the number of display impression as well as the number of sites and pages containing the display ads all have a positive impact on the repeat purchase probability. A recent ComScore report also finds the banner ads impression could be more influential in leading to conversions than the click-throughs (Lipsman 2012). The carryover effects of Organic Search, Paid Search and Referral are also significantly positive. This implies that for the focal firm more repeated visits to the website through these channels are indicative of the greater attractiveness of the firms' offering vis-à-vis their competitors and thus indicative of a higher likelihood of purchase. The carryover effect of Direct visits is also positive, consistent with Bowman and Narayandas (2001)'s finding that customers who directly visit the firm's site more often may have a stronger preference for the firm's offering and thus leading to a positive carryover. All of these results strongly support Hypothesis 3.

With regard to spillover, we find informational stock of Organic Search has a positive spillover on purchases through Paid Search channel, while the reverse effect is not significant. While informational stock of Display has a positive spillover on purchases through E-Mail channel, the reverse spillover is not significant. The spillover effects of informational stock of firm-initiated channels are, by and large, positive on purchases through customer-initiated channels, except for the effect of informational stock of Display on Referral channel which is significantly negative, partially supporting Hypothesis 4. This may indicate that customers who visit through Display click-through often may use the Referral channel for gathering additional information but may not consummate purchase through that channel. It is also interesting to note that the spillover of informational stock of Organic and Paid Search are all negative (when significant) on purchases through Referral, E-Mail and Display channels. Given that, at the visit stage, the spillovers of Search channels contribute to reducing the costs of visiting in Referral, E-Mail and Display channels, one can similarly surmise that the customers who visit the website through search channels often use these other channels mainly for gathering information but not for making purchases on those visits. In short, search can help in bringing in more visits, but not necessarily more conversions. Additionally, the spillover of other channels on Paid Search and Direct purchases are always positive indicating that the informational stock of other channel visits help conversion during Paid Search and Direct visits.

Overall, our results are consistent with our conceptual model and lend good support to the hypotheses we posit. There are significant carryover and spillover effects both at visit and purchase stages.

The estimated decay rates of information gathered in a channel provide insights into how fast the informational stock accumulates in each channel. We observe that the decay rates are generally low for the Search channels and E-Mail channel (.27 for Organic Search, .38 for Paid Search, and .31 for E-Mail), while it is the highest for Display channel (.53). Thus, a search click-through or an e-mail click-through has significantly long lasting impact, while a Display impression or click-through has the least enduring impact. Viewing this from a complementary perspective, display retains only .5% of its original informational value after 7 days, while Organic Search retains 11.0%, Paid Search 3.5% and E-mail 7.4%. The corresponding values for Referral and Direct are in the 2% range. While the informational value of an E-mail is understandable given that it can be retrieved and used again, the informational value of searches also having longer lasting impact is an interesting and useful finding.

Next, we account for these carryovers and spillovers in estimating the contribution of the different channel visits to the overall conversion to get a better picture of the relative contributions of the channels than what a "last-click" model can provide us.

Estimating contribution to purchases

Given the calibration data and the estimates from Table 3, we estimate the impact of a specific channel, say e-mail, on predicted probabilities of conversion by excluding e-mail from the proposed model to predict the probabilities of conversion without e-mails. The difference between the predicted number of conversions with and without e-mails should provide us an estimate of the incremental value of e-mails in the calibration data in affecting conversions through e-mail channel as well as other channels. However, the above estimates are incremental, given that other variables (channels) already exist in the model, and may already explain significant variance in the dependent variable. Therefore, using the idea of Shapley value in game theory (Shapley 1953), we calculate the total contribution of each channel in leading to a conversion by averaging over their incremental contributions in all possible channel combinations formed at the visit stage. Specifically, we adopt an approach similar to Kruskal's "relative importance" theory (Kruskal 1987) and add each channel stepwise to estimate the predicted probabilities of conversion and measure the increase in predicted total conversions, denoted by δ. Since the channel entering the model earlier may explain a larger bulk of the predicted purchase than those entering later, we explore all the possible entering orders $\{k\}$ and take the average over all the increased predictions δ_k associated with each channel (Menard 2004). Based on this analysis, the last 2 columns in Table 4 (Tables follow Reference throughout) shows the contribution of each channel to purchase conversions, which is compared against the two widely-used metrics in the industry: (1) the last-click attribution metric which gives the entire credit to the visit when conversion occurred and (2) 7-day average attribution metric which assigns the conversion credit equally to all the visits made in the past 7 days. Note that these metrics, unlike our model, use touch data ended in conversions and exclude all non-conversion data.

While attribution percentages across channels differ between Last-Click and 7-Day Average models their conversion ranks stay the same in both models. However, our proposed model alters the attribution percentages and ranks significantly by accounting for the carryovers and spillovers. For example, the attribution of Organic Search drops significantly from 25% to 16%, while Paid Search decreases to 6% and drops to the last rank. While Referral channel climbs to the second rank with 24%, E-Mail and Display attributions almost double their number of conversions credited in Last-Click model. Our result show that there are significant changes in attributions which would have far-reaching implications for ROI and budget allocations for marketing interventions such as paid search, display and e-mail. In Table 3 all other channels have positive spillovers in enabling purchases through the Direct channel, which could account for the drop in its attribution, although Direct also gains from spillovers to other channel. The most dramatic drop in attribution is in Organic Search, which has positive spillover from Referral and E-Mail, both of which gain in attribution probably at the expense of Organic Search. These results clearly highlight the importance of considering the path data in estimating attributions of the channels and accounting for the carryover and spillover effects across channels on conversion. This also suggests the firm could intervene with marketing actions that could possibly play a positive role in effecting conversions at the website, which is discussed in the following subsection.

Path sequence and marketing interventions

A key insight that emerges from our results is the understanding of whether and when to intervene with marketing actions given a customer's path to the firm's website. Since the model provides us the estimates of the impact of previous visits (the lag estimates in Table 3), it is possible to predict for a customer, given his/her path data to date, the probabilities of visit through different channels for the next visit occasion and the probability of a purchase on that visit under different intervention scenarios. We illustrate this with an example of e-mail intervention. In our calibration sample, e-mail interventions target a significant number of customers regardless of their rewards program status – specifically, 23% of the non-members and 45% of the members were targeted, with the content of the e-mail the same across customers. To stay within the confines of the calibration model for our illustration, we focus our analysis only on customers who have already been targeted with e-mail interventions in their past. Thus, our objective is to understand under what path characteristics the firm can increase the overall probability of conversion for a customer who has had a prior e-mail intervention in his/her path by targeting the customer with another e-mail intervention; and under what conditions the firm is better off not targeting them by another e-mail.

Table 5 (Tables follow Reference throughout) provides these probability estimates for selected instances of path data that have prior e-mail interventions. In Row 1, a customer is observed for the first-time entering the website on Day (T-2) through Organic Search channel, makes another visit through E-Mail channel on Day (T-1). If there is no intervention, the total probability of purchase through any channel on Day T is .447, with a visit most likely through Organic Search. However, an E-Mail intervention on Day T increased the total probability of purchase to .474. Table 5 provides many such scenarios. It is seen that when a visit on Day (T-1) happens through the Direct channel (Rows 9 and 12), the best option for the firm is to not intervene as E-mail intervention can only lower the likelihood of conversion. Rows 13 through 20 provide similar scenarios where the advantage of e-mail targeting is clearly contingent upon the path taken by a customer. This illustration provides the utility of our approach for retargeting customers with marketing interventions. If customers' history of path sequences is tracked once they enter the website for the first time, the firm can use the data to customize the marketing interventions for each identified customers to maximize their purchase probability. For a full-fledged implementation of such individualized targeting, the criterion used for targeting, especially in display channel, has to be worked into a supply side equation. Also, using a dynamic optimization procedure (Li, Sun, and Montgomery 2011) a firm can identify optimal targeting policies considering customers' current and future probabilities of purchase.

Field study with paid search off

Our model helps managers in understanding the incremental effect of each channel and predicting their impact on conversions. Even in situations when one channel (say, Paid Search) were to be turned off, our model is able to predict the reallocation of channel shares in leading to conversions. To test and further validate our model, we obtained a validation sample covering the period August through November, where for one week, Nov 3 through Nov 9, the firm shut down the Paid Search option completely. Using this validation sample, we made two sets of predictions of conversions for this one week period when Paid Search was off. The first set of predictions (Paid Search Off) was based on the fact that Paid Search channel was not available for any customers to consider or choose. Since we have explicitly modeled the consideration set of consumers, we can constrain consideration probabilities of Paid Search channel to be zero in estimating this set of predictions. The second set of predictions (Paid Search On) was made by assuming that all channels were available for this one week (note that our model was calibrated on a sample with all channels available).

Table 6 (Tables follow Reference throughout) provides the two sets of predicted conversions along with the observed conversions during this week. First, in comparing the total predictions with Paid Search On and Paid Search Off, we find that overall conversions drop from 11,893 to 11,106, a decrease of 6.6% in conversions. This drop could be due to the absence of Paid Search – that is, the incremental contribution of Paid Search for this sample, which is lost when Paid Search is turned off. This is less than the 923 conversions (7.8% of total conversions) predicted for the Paid Search channel when assuming all channels are available. It appears that some of the Paid Search conversions are being recaptured by other channels when Paid Search is turned off (see Column 3) resulting in only a 6.6% drop in conversions rather than the 7.8% or more.

Second, the prediction for total conversions with Paid Search Off (11,106) is very close to the observed conversions observed in the study (11,395) with a MAPE of 2.6%. This validates the ability of our model in predicting conversions when a specific channel is not available, and illustrates how our model can be used to estimate the incremental contribution of a channel. Third, comparing the predicted conversions with Paid Search Off and the observed conversions channel by channel, we find that the observed conversions through Organic Search is much higher (MAPE=30%), with Referral conversions also being higher (MAPE=21%) while Direct conversions are lower (MAPE=16%) than what our model predicted. When Paid Search is shut off, the prediction intervals based on the 95% highest posterior density (HPD) covers the observed conversions for all channels except Organic search. This shows the strength of our model in prediction even when one channel is shut off. In fact, our model performs much better than a model that does not take the consideration stage into account. We further investigated the prediction variance of Organic Search, by segmenting the Paid Search conversions in the validation sample with "branded" and "unbranded" keywords. Approximately 73% of the Paid Search conversions are based on "branded" keywords, while the rest (27%) are through "unbranded" keywords. Since the firm has a very strong brand, their relative rank of branded keywords in the Organic Search pages is almost always the first, while for many unbranded keywords they bid on, the firm also ranks within the first webpage of Organic Search results. Thus, when Paid Search is off, it appears that much of the conversions previously stemming from paid branded keywords are being recaptured by free Organic Search, instead of being "lost", while a good percentage of "unbranded" keyword conversions do get lost. This could possibly explain why the observed conversions through Organic Search is much higher (43%) than what the model predicted, and the observed overall conversions is somewhat higher (3%) than what the model predicted. In sum, given the firm's brand strength and 73/27 split between branded and unbranded keyword in Paid Searches, the recapture rate of Paid Search conversions when pausing Paid Search is higher than what the model predicts.

Conclusions

We have proposed a conceptual framework to measure and estimate the carryover and spillover effects across online marketing channels through which customers visit a firm's website. The model forms the basis for attributing and allocating credit for conversions to both marketing-initiated and customer-initiated channels. Ours is the first study, to our knowledge, which examines these effects in the online channel context at the distinctly different stages – visit and purchase. Our empirical study illustrates the importance of estimating these effects so that the attribution of each channel to the overall conversions at the website can be accurately determined. This has useful managerial implications for allocating marketing budget across marketing channels and for targeting strategies. We will first examine the implications for the specific context we have studied, and then discuss the more general implications.

Implications for the focal firm

Our study finds significant spillover effects of firm-initiated channels to customer-initiated channels both at the visit stage and at the purchase stage. Firm-initiated interventions also impact visits in the short-term with no long-term carryover effects. This implies that managers have to take a more inclusive and macro view of the returns to investments in firm-initiated interactions. All impact considered, the contribution of E-Mail and Display ads to conversions is significantly underestimated. Similarly, the role of Referral channel is also underestimated by the last-click model. Significantly, the real impact of Organic Search on conversions is much lower than what it appears to be in the last-click model. For the focal firm it is clear that some customers, having visited the website through other channels previously, are using Organic Search purely as a navigational tool to get to the website in completing purchases. The impact of Paid Search and Direct are also diminished. Given that the changes in attributions based on our proposed model are considerably different (ranging from -40% to +75%), it clearly implies a different allocation of marketing budget. The focal firm in our study uses the attribution estimates to charge their franchisees for the various marketing programs such as paid search, referrals, and other campaigns, so even if the attribution ranks were only marginally different it would still make a sizable difference for such appropriations. Attributions based on our model would render these appropriations in line with the incremental purchases that the franchisees actually observe at their properties. This will enhance franchisees' confidence in such metrics and the fairness perception of the firm in how they pass on the marketing costs. Our attribution model is designed to be estimated and run for each period, say a month, so that it becomes the basis for allocating the marketing expenses and attribution for each channel for each month. This can also form the basis for determining the acquisition costs through each channel and understand the efficacies of each channel in each period.

While our results show that E-Mail and Display ads are effective in the short-run, it is important that they are not used indiscriminately to target all visitors to the websites using the often-used strategy of "retargeting", where e-mails and displays follow visitors everywhere once they click on an e-mail, display ad or visit the website (Helft and Vega 2010). As our path analysis results show, retargeting visitors to the website with e-mails is not always the best strategy. While in some cases e-mail retargeting increases the overall purchase probability for those customers, in other cases it actually hurts the purchase probability for the same segment of customers. Our model can be used for such customized targeting, and pinpoint the cases where they are likely to contribute to higher conversions.

Finally, our model allows us to estimate how conversions through different channels are affected when a channel is not available. We observe that a significant portion of the conversions that could have occured through Paid Search channel is recaptured through Organic Search. Since the firm in our example has a strong brand and ranks very high in Organic Search, we conjecture that many of the branded keyword searches that could have occurred through Paid

Search are recaptured through Organic Search. Thus, the incremental contribution of Paid Search to conversions is much lower than what a last-click model would lead us to believe, and the firm can reallocate marketing investments given this knowledge.

General implications

It is clear from our study that the last-click attribution model or the 7-day average model are not good metrics for understanding the real impact of firm-initiated channels as well as customer-initiated channels on conversions. These metrics consider only those visits that result in conversion immediately to the exclusion of the other visits. While they may provide passable results in product categories with very short purchase funnel (with one or two touch points) and with fewer channels, they will invariably be misleading in product/service categories with longer purchase funnels as in high involvement categories – consumer durables and travel services – and for firms with multiple channels, both customer-initiated and firm-initiated. In the latter case, we also expect that the effectiveness of firm-initiated efforts will be underestimated using the last-click model and it is imperative that firms use our framework to estimate the real incremental impact. In addition, there are limits to the effectiveness of firm-initiated efforts such as e-mail and display retargeting campaigns. Recent reports (Mattioli 2012) suggest that retailers are finding that overuse of e-mails actually annoys many customers and becomes less effective. Our path analysis confirms that repeated e-mail targeting works only in certain cases and not across all cases and retailers may be well-advised to carefully calibrate the targeting of customers with these instruments.

Our results suggest that the incremental impact of Paid Search channel may not be as high as what the last-click model suggests, and if Paid Search were to be discontinued, much of its impact can be recaptured through the Organic Search channel. The generalizability of this result, however, depends on the brand strength of the firm. If the brand is not very strong, then such recaptures may not materialize as the firm may not get a high enough position in Organic Search. We conjecture that the stronger the brand, the lower the incremental effect of Paid Search on ultimate conversion. Our framework provides a useful tool to determine this incremental contribution and to determine if the cost of effecting a conversion through Paid Search is less than the incremental revenue obtained through the channel. Since paid search makes up 59% of the overall spending in online marketing budget for many firms (VanBoskirk et al. 2009), such analysis can be useful to contain marketing costs through very selective use of keywords and possible negotiations with search engine companies.

One of the useful features of our model is that it incorporates customers' consideration sets of channels to use in visiting the firm's website. As there is significant heterogeneity and self-selection in customers' consideration of channels to use, by modeling consideration sets endogenously, our modeling framework allows us to accurately predict the conversions through different channels when one of them (for example, paid search, as in our study) is not available.

Limitations and future research

Given that our model is estimated using secondary data and not experimental data, there is a possibility that alternative explanations exist for the effectiveness of display and e-mail campaigns such as selective targeting of customers with inherently higher propensity to purchase (Manchanda, Rossi, and Chintagunta 2004). This problem is somewhat mitigated in our study as e-mail is not specifically targeted – e-mail offers are not just sent to rewards program members, but also to all past purchasers and all visitors with e-mail registration irrespective of which channel they usually visit. With respect to display, targeting is an issue as the firm uses Doubleclick as a vendor. To check whether such targeting is correlated with the channels customers often use or with their rewards program membership, we estimated the incidence of display impressions and conversions across customers visits through different channels, and as well as across non-members and rewards levels. A similar exercise was conducted with e-mail incidence and conversions. Both analyses revealed that the correlations were minimal, indicating there is no systematic pattern in targeting, at least not on the observed dimensions of channels and rewards program membership. Although our results are conditional on firm's ongoing targeting strategies, we believe that the effects of strategic targeting are not likely to change the essential nature of our results. The focal firm provides a variety of substitutable products in a wide price range. Customers with different budgets can easily find their affordable choice within the target firm. To minimize selectivity bias, we can compare the results of our analyses on different cohorts of visitors separated a spell of one month or more, and use the observed variations in the firm's targeting and promotional campaigns to make the results more useful.

We have currently modeled customer visits using a static framework. However, in the context of planned purchases, customer visits could be modeled in a dynamic setting taking into account their forward-looking and strategic behavior. Long term dynamic changes in search behavior and purchase decisions can be examined using structural models with appropriate long term data. This could be a possible extension to our framework. We do not model the supply-side decision of the firm in targeting customers for e-mail campaigns, selecting locations for banner ads, and the keywords to bid on for paid search, yet the data of the conversion path are conditional on the above decisions that the firm has already made. Given this endogeneity, our model measures the relative effectiveness of these channels conditional on the decisions made by the firm. To examine the impact of marketing interventions under policies very different from the one being used, modeling supply-side decisions would be very useful. We leave this for future research.

Finally, the model we have developed has a broader application beyond the B2C context. For example, in business markets, sales conversion is often preceded by multiple vehicles of marketing effort such as trade shows, direct mails and e-mail campaigns, salesperson visits and so on, and our framework and methodology should be well suited to analyze such contexts.

References

Ansari, Asim and Carl F. Mela (2003), "E-customization," *Journal of Marketing Research*, 40 (2), 131-45.

Ansari, Asim, Carl F. Mela, and Scott A. Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research*, 45 (1), 60-76.

Atlas (2008), "Engagement Mapping - A New Measurement Standard Is Emerging for Advertisers," *Atlas Institute's Digital Marketing Insights*.

Bowman, Douglas and Das Narayandas (2001), "Managing Customer-Initiated Contacts with Manufacturers: The Impact on Share of Category Requirements and Word-of-Mouth Behavior," *Journal of Marketing Research*, 38 (3), 281-97.

Bucklin, Randolph and Catarina Sismeiro (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, 40 (3), 249-67.

Chan, Tat Y., Chunhua Wu, and Ying Xie (2011), "Measuring the Lifetime Value of Customers Acquired from Google Search Advertising," *Marketing Science*, 30 (5), 837-50.

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

Chib, Siddhartha and Edward Greenberg (1995), "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49 (4), 327-35.

Danaher, Peter J., Guy W. Mullarkey, and Skander Essegaier (2006), "Factors Affecting Web Site Visit Duration: A Cross-Domain Analysis," *Journal of Marketing Research*, 43 (2), 182-94.

De Los Santos, Babur, Hortaçsu, Ali, and Wildenbeest, Matthijs R. (2012), "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," American Economic Review, 102 (6), 2955-80.

DoubleClick (2004), "Retail Details: Best Practices in Multi-Channel Integration," *A J.C. Williams Group Report*, March 2004.

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

Erdem, Tülin and Michael P. Keane (1996), "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15 (1), 1-20.

Gelman, Andrew, Xiao-Li Meng, and Hal Stern (1996), "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies," *Statistica Sinica*, 6 (4), 733-807.

Geweke, John (1992), "Evaluating the Accuracy of Sampling- Based Approaches to the Calculation of Posterior Moments," in *Bayesian Statistics 4*, J. M. Bernardo, J. O. Berger, A. P. Dawid, A. F. M. Smith, eds. Oxford: Oxford University Press.

Ghose, Anindya and Sha Yang (2009), "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets," *Management Science*, 55 (10), 1605-22.

Gilbride, Timothy J., and Peter J. Lenk (2010), "Posterior Predictive Model Checking: An Application to Multivariate Normal Heterogeneity," *Journal of Marketing Research*, 47 (5), 896-909.

Green, Cindy Estis (2008), "Demystifying Distribution 2.0," *TIG Global Special Report*, McLean: The Hospitality Sales and Marketing Association International Foundation.

Hauser, John R. and Birger Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16 (4), 393-408.

Helft, Miguel and Tanzina Vega (2010), "Retargeting Ads Follow Surfers to Other Sites," *The New York Times*, August 29, 2010.

Ilfeld, Johanna S. and Russell S. Winer (2002), "Generating Website Traffic," *Journal of Advertising Research*, 42 (5), 49-61.

Johnson, Eric J., Steven Bellman, and Gerald L. Lohse (2003), "Cognitive Lock-In and the Power Law of Practice," *Journal of Marketing*, 67 (2), 62-75.

Keane, Michael P. (1992), "A Note on Identification in the Multinomial Probit Model," *Journal of Business and Economic Statistics*, 10 (2), 193-200.

Kim, Jun B., Paulo Albuquerque, and Bart J. Bronnenberg (2010), "Online Demand under Limited Consumer Search," *Marketing Science*, 29 (6), 1001-23.

Kruskal, William (1987), "Relative Importance by Averaging Over Orderings," *The American Statistician*, 41 (1), 6-10.

Kumar, V. and Rajkumar Venkatesan (2005), "Who Are the Multichannel Shoppers and How Do They Perform?: Correlates of Multichannel Shopping Behavior," *Journal of Interactive Marketing*, 19 (2), 44-62.

Kumar, V., Vikram Bhaskaran, Rohan Mirchandani, and Milap Shah, “Creating a Measurable Social Media Marketing Strategy for Hokey Pokey: Increasing the Value and ROI of Intangibles and Tangibles,” forthcoming, Marketing Science.

Latham, Steve (2011), "Attribution 101: Full-Funnel Media Measurement," *Encore Media Metrics*.

Li, Shibo, Baohong Sun, and Alan L. Montgomery (2011), "Cross-Selling the Right Product to the Right Customer at the Right Time," *Journal of Marketing Research*, 48 (4), 683-700.

Lipsman, Andrew (2012), "For Display Ads, Being Seen Matters More than Being Clicked," *comScore Press Release*, April 24, 2012.

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

Manchanda, Puneet, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K. Chintagunta (2006), "The Effect of Banner Advertising on Internet Purchasing," *Journal of Marketing Research*, 43 (1), 98-108.

Martin, Andrew (2009), "The Long Road to Conversion: The Digital Purchase Funnel," *Digital Marketing Insight*, Microsoft's Atlas Institute.

Mattioli, Dana (2012), "Stores Smarten Up Amid Spam Flood," *The Wall Street Journal*, March 9, 2012.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2003), "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science*, 22 (1), 58-84.

Menard, Scott (2004), "Six Approaches to Calculating Standardized Logistic Regression Coefficients," *The American Statistician*, 58 (3), 218-23.

Moe, Wendy W. and Peter S. Fader (2004), "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science*, 50 (3), 326-35.

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, and John C. Liechty (2004), "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23 (4), 579-95.

Moorthy, Sridhar, Brian. T. Ratchford, and Debabrata Talukdar (1997), "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Research*, 23 (4), 263-77.

Mulpuru, Sucharita, Shar VanBoskirk, Patti Freeman Evans, and Doug Roberge (2011), "The Purchase Path of Online Buyers," *Forrester Report*.

Myers, Joseph B., Andrew D. Pickersgill, and Evan S. Van Metre (2004), "Steering customers to the right channels," *The McKinsey Quarterly*, November 2004.

Naik, Prasad A. and Kalyan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communications," *Journal of Marketing Research*, 40 (4), 375-88.

Neslin, Scott A. and Venkatesh Shankar (2009), "Key Issues in Multichannel Customer Management: Current Knowledge and Future Directions," *Journal of Interactive Marketing*, 23 (1), 70-81.

Oliver J. Rutz, Michael Trusov and Randolph E. Bucklin (2011), "Modeling Indirect Effects of Paid Search Advertising: Which Keywords Lead to More Future Visits?," *Marketing Science*, 30 (4), 646-65.

Punj, Girish and Richard Brookes (2002), "The Influence of Pre-Decisional Constraints on Information Search and Consideration Set Formation in New Automobile Purchases," *International Journal of Research in Marketing*, 19 (4), 383-400.

Rutz, Oliver J. and Randolph E. Bucklin (2011), "From Generic to Branded: A Model of Spillover in Paid Search Advertising," *Journal of Marketing Research*, 48 (1), 87-102.

Shapley, Lloyd S., (1953) "A Value for n-person Games," in Contributions to the Theory of Games, volume II, by H.W. Kuhn and A.W. Tucker, editors, *Annals of Mathematical Studies*, v. 28, pp. 307-17. Princeton University Press.

Sherman, Lee and John Deighton (2001), "Banner Advertising: Measuring Effectiveness and Optimizing Placement," *Journal of Interactive Marketing*, 15 (2), 60-64.

Shugan, Steven M. (1980), "The Cost of Thinking," *Journal of Consumer Research*, 7 (2), 99-111.

Song, Jaeki and Fatemeh "Mariam" Zahedi (2005), "A Theoretical Approach to Web Design in E-Commerce: A Belief Reinforcement Model," *Management Science*, 51(8), 1219-35.

Stephen, Andrew T. and Jeff Galak (2012), "The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace," *Journal of Marketing Research*, 49 (5), 624-39.

Terui, Nobuhiko, Masataka Ban, and Greg M. Allenby (2011), "The Effect of Media Advertising on Brand Consideration and Choice," *Marketing Science*, 30 (1), 74-91.

Valentini, Sara, Elisa Montaguti, and Scott A. Neslin (2011), "Decision Process Evolution in Customer Channel Choice," *Journal of Marketing*, 75 (6), 72-86.

VanBoskir, Shar, Christine Spivey Overby, Jennifer Joseph McGann, Niki Scevak, and Angie Polanco (2009), "US Interactive Marketing Forecast, 2009 to 2014," *Forrester Report*.

van Nierop, Erjen, Bart Bronnenberg, Richard Paap, Michel Wedel, and Philip Hans Franses (2010), "Retrieving Unobserved Consideration Sets from Household Panel Data," *Journal of Marketing Research*, 47 (1), 63-74.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (4), 106-25.

Wiesel, Thorsten, Koen Pauwels, and Joep Arts (2011), "Marketing's Profit Impact: Quantifying Online and Off-line Funnel Progression," *Marketing Science*, 30 (4), 604-11.

Yang, Sha and Anindya Ghose (2010), "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?" *Marketing Science*, 29 (4), 602-23.

Zauberman, Gal (2003), "The Intertemporal Dynamics of Consumer Lock-In," *Journal of Consumer Research*, 30 (3), 405-19.

Tables

Table 1: Summary Statistics

Channel	Channel Visits	Purchases	Purchase/Visit Conversion Rate
Organic Search	4469	285	6.38%
Paid Search	1557	114	7.32%
Referral	3980	201	5.05%
Direct	7959	347	4.36%
E-Mail	2804	138	4.92%
Display	1600	43	2.69%
Total	**22369**	**1128**	**5.04%**

Table 2: Model Comparison

Channel	Model 1	Model 2	Model 3	Model 4	Proposed Model
Organic Search	0%	10%	43%	237%	30%
Paid Search	6%	19%	90%	190%	3%
Referral	120%	14%	193%	311%	21%
Direct	124%	30%	71%	863%	14%
E-Mail	98%	23%	29%	189%	15%
Display	84%	37%	71%	2076%	33%
Overall	74%	20%	35%	502%	1%
Log-Marginal Likelihood	-13580	-14692	-17033	-173093	-12521

Notes: All the percentage values in this table are mean absolute percentage errors (MAPE)

Table 3: Model Estimates

Channels ***Variables***	**Organic Search**	**Paid Search**	**Referral**	**Direct**	**E-Mail**	**Display**
		(Estimates are posterior means)				
Consideration Stage:						
Intercept	**1.60**	**1.84**	**2.43**	**2.65**		
Rewards Level-1	.04	.04	**.92**	**.59**		
Rewards Level-2	-.03	**-.15**	**.74**	**.69**		
Rewards Level-3	**-.16**	**-.18**	**.46**	**1.92**		
Rewards Level-4	**-.17**	**-.19**	**1.00**	**.94**		
Visit Stage:						
Intercept	**2.27**	**1.26**	**-.92**	**.40**	**-.36**	**1.92**
Search Cost	**-1.37**	**-1.96**	**-3.58**	**-3.11**	**-3.58**	**-1.56**
τ (tau)	**.35**					
Cost:						
Cumulative time	-.77	**-1.15**	**-.99**	**-1.41**	-.78	-.79
Lag Organic Search	**-2.10**	**-.18**	**-.20**	.07	**-.30**	**-.25**
Lag Paid Search	**-.79**	**-1.97**	**-.19**	.11	**-.49**	**-.43**
Lag Referral	**-.38**	**-.13**	**-2.43**	.05	.12	.01
Lag Direct	**.47**	**-.29**	.03	**-1.71**	**.19**	-.01
Lag E-Mail	**.74**	**-.18**	**-.21**	**.24**	**-2.04**	**.49**
Lag Display	**-.27**	**-.27**	**.16**	-.04	.11	**-1.26**
Lag No Visit	1	1	1	1	1	1
Purchase Stage:						
Intercept	**-1.29**	**-.94**	**-1.11**	**-1.29**	**-1.38**	**-1.39**
Info Stock - Organic Search	**.68**	**.17**	**-.39**	**.21**	**-.21**	-.12
Info Stock - Paid Search	.03	**.44**	.03	**.23**	.04	**-.26**
Info Stock - Referral	**.16**	.03	**.35**	**.18**	.11	**.44**
Info Stock - Direct	-.11	**.22**	**.70**	**.73**	**.22**	**.47**
Info Stock - E-Mail	**.28**	**.61**	-.15	.08	**.83**	.06
Info Stock - Display	.07	**.16**	**-.38**	**.22**	**.28**	**.40**
λ=(1- Decay Rate)	**.73**	**.62**	**.57**	**.59**	**.69**	**.47**

Notes: Bold indicates that the 95% posterior interval excludes zero.
τ is the coefficient of the inclusive value.

Table 4: Contribution to Conversions

Channel	Observed	Last Click %	Last Click Ranking	7-day Average %	7-day Average Ranking	Proposed Model %	Proposed Model Ranking
Organic Search	285	25%	2	24%	2	16%	4
Paid Search	114	10%	5	8%	5	6%	6
Referral	201	18%	3	18%	3	24%	2
Direct	347	31%	1	30%	1	28%	1
E-Mail	138	12%	4	14%	4	19%	3
Display	43	4%	6	6%	6	7%	5
Total	**1128**	**100%**		**100%**		**100%**	

Table 5: Path Sequence and Visit/Purchase Probabilities

	Day(T-4)	Day (T-3)	Day (T-2)	Day (T-1)	Day T				
					No intervention			E-mail intervention	
Row	Visit thru	Visit thru	Visit thru	Visit thru	Visit Prob.	Purchase Prob.	Visit thru	Click Prob.	Purchase Prob.
1	X	X	OS	E-Mail	.196	.447	OS	.185	**.474**
2	X	X	PS	E-Mail	.193	.446	OS	.182	**.473**
3	X	X	R	E-Mail	.193	.447	OS	.182	**.474**
4	X	X	D	E-Mail	.190	.459	D	.179	**.495**
5	X	X	E-Mail	E-Mail	.194	.450	OS	.183	**.476**
6	X	X	E-Mail	OS	.295	.317	OS	.193	**.466**
7	X	X	E-Mail	PS	.298	.324	PS	.182	**.465**
8	X	X	E-Mail	R	.238	.343	R	.181	**.467**
9	X	X	E-Mail	D	.421	.565	D	.208	.512
10	X	E-Mail	OS	PS	.230	.356	PS	.184	**.463**
11	X	E-Mail	OS	R	.238	.341	R	.184	**.465**
12	X	E-Mail	OS	D	.421	.564	D	.208	.510
13	E-Mail	OS	X	X	.290	.166	OS	.188	.150
14	E-Mail	PS	X	X	.287	.136	PS	.188	**.150**
15	E-Mail	R	X	X	.214	.137	OS	.188	**.150**
16	E-Mail	D	X	X	.359	.335	D	.187	.149
17	OS	E-Mail	X	X	.180	.172	OS	.217	**.191**
18	PS	E-Mail	X	X	.190	.136	PS	.216	.121
19	R	E-Mail	X	X	.222	.196	OS	.252	.190
20	D	E-Mail	X	X	.220	.194	OS	.215	.189

Notes: OS is Organic Search; PS is Paid Search; R is Referral; D is Direct; X is no visit.
Bold number in the last column indicates the purchase probability is increased with e-mail intervention.

Table 6: Predicted Conversions – Field Study

Channel	Observed Conversions	Paid Search Off			Assuming Paid Search On	
		Predicted	95% HPD region	MAPE	Predicted	95% HPD region
Organic Search	2453	1711	[1553 1854]	30%	1023	[869 1192]
Paid Search	0	0			923	[782 1071]
Referral	2271	1784	[1376 2308]	21%	2775	[2231 3226]
Direct	5398	6269	[4320 7426]	16%	5785	[4204 7410]
E-Mail	1114	1260	[1109 1349]	13%	907	[782 1049]
Display	159	82	[19 207]	48%	480	[378 569]
Total	11395	11106		2.60%	11893	

Figure

Figure 1: Conceptual Framework