Multichannel Path-to-Purchase: Channels as "Engagers" and "Closers"

Marcel Goić
Kinshuk Jerath
Kirthi Kalyanam

August 2017

Preliminary draft, comments welcome

Marcel Goić is Assistant Professor of Marketing at the Department of Industrial Engineering, University of Chile, Republica 701, Santiago, Chile (phone:562-29784810, email: mgoic@dii.uchile.cl). Kinshuk Jerath is Associate Professor of Marketing at Columbia Business School, 3022 Broadway, New York, NY 10027 (phone:12128542294, email: jerath@columbia.edu). Kirthi Kalyanam is the J.C. Penney Research Professor and the Director of the Retail Management Institute at the Leavey School of Business at Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053 (phone: 408-554-2705, email: kkalyanam@scu.edu).

The first author gratefully acknowledges partial funding by Fondecyt (Project No 11121434) and the Millennium Institute on Complex Engineering Systems.

Multichannel Path-To-Purchase: Channels as "Engagers" and "Closers"

August 2017

Abstract

In today's retailing environment, consumers and sellers interact through multiple channels (e.g., online channels include email, search engines, banner ads, affiliate websites, comparison shopping websites, etc.) and it is critical for a seller to understand the roles played by different touchpoints on a consumer's "path to purchase." To this end, we propose a parsimonious model in which the propensity to visit and purchase using each channel is dependent on latent "inventories of experience" that accumulate based on the consumer's path up to that point. We allow for clustered/clumpy visits by consumers which enables separation of their behavioral patterns into intra- and inter-session components. We estimate the model on a dataset obtained from an online retailer and find that different channels play different roles in the purchase process. Channels may act as "closers" or "engagers"—customers arriving through the former are more likely to make a purchase, while customers arriving through the latter, even if they don't make a purchase, are more likely to visit again in the future. We also find that the role played by a channel depends on the browsing history up to that point and may change over time. Furthermore, we show that some channels can have a negative impact on future activity. Finally, we show that the path that a customer has taken until a particular point in time has high predictive power for her future visit and purchase activity.

Keywords: path-to-purchase, online marketing, probability models.

1. Introduction

Firms employ multiple channels to contact customers, ranging from traditional media advertising to newer Internet marketing tools. For instance, an online retailing firm can attract customers to and motivate purchases on its website by using several promotional vehicles such as search engines, emails, price comparison websites and links at affiliate websites.[1] In this setting, it is critical for a retailer to understand the role played by each channel in the "path to purchase" of a consumer. For instance, some channels may help to keep a relationship active with a customer, others may help to monetize the transactions between the firm and the customer. Furthermore, there may be interdependencies among the various channels, which may depend on the order in which the customer was exposed to the different channels, and the assessment of marketing interventions must carefully account for these interdependencies. An understanding of this is a key component in the design of a successful marketing campaign and in the allocation of resources to different channels.

In this research, we posit that browsing and purchase decisions are affected by the full set of contacts that customers have with the company, i.e., the customer's browsing path. We propose a probability model for a multichannel marketing campaign that summarizes the impact of the browsing path of a customer using a time-variant latent variable, which we call the "inventory of experience" of the customer with the firm. The observed behavior of the customer is stochastically dependent on this latent variable. Our model therefore captures the impact of the browsing path of the customer on her likelihood of engaging with the firm in the future, as well as on the likelihood of converting, i.e., our model can examine both the immediate and future impact of a channel. Notably, we allow a channel to have a negative effect on future engagement and conversion

[1] Our empirical setting is that of an online retailer. The different channels for our case are described in Table 1.

probabilities. Our framework helps to understand the role and relative importance of each channel in impacting a customer's behavior. Using our model, we can also predict future browsing and purchase activity at individual and collective levels.

We estimate our model on data obtained from an online retailer selling a durable good. Customers visit this retailer's website through a number of different channels and we find that different channels indeed play different, and complementary roles in influencing customers. Some channels have more short-term value as "closers," i.e., customers arriving through these channels are more likely to make a purchase in the current visit; these include arriving at the website by clicking on a result at a search engine after searching for a keyword containing the retailer's brand name, or arriving by directly typing the URL in address bar of the browser. Other channels have longer-term value as "engagers," i.e., customers arriving through these channels, even if they don't make a purchase in the current visit, are more likely to visit again in the future; these include arriving at the website by clicking on an email communication sent by the firm. Furthermore, arrival through some channels indicates a reduced likelihood of making a purchase. Typically, these are channels that provide information on competitors; for instance, arriving by clicking on a result at a search engine after searching for a keyword that is related to the category but does not contain the retailer's brand name, and arriving by clicking on a search result at a shopping engine (e.g., a price comparison engine). Therefore, channels can also be classified as "competitive" (i.e., arrival through these channels indicates that the customer may be considering competitors) versus "captive" (i.e., arrival through these channels indicates high interest in the company). We also find that the browsing history of a customer has an important role to play and ignoring this history can generate incomplete views on customer behavior and incorrectly estimate the channels' abilities to generate value to the company. Moreover, we show that the path that a customer has taken until a particular point in time has high predictive power for her future visit and purchase activity.

There is a nascent literature addressing the impact of path to purchase using different modeling approaches. Abhishek et al. (2014) develop a Hidden Markov Model to capture state and path dependency in customers' conversions. Li and Kannan (2014) develop a multilevel hierarchical model to capture customers' stages leading to a conversion. Anderl et al. (2013) use graph-based models to estimate the marginal impact of different channels on customers' decisions. Xu et al. (2014) use a mutually exciting point process model for attributing a conversion to a limited number of channels on the path of a customer. The focus of these papers is attribution of outcomes to different channels, which is different from our objective which is that given a certain history of contacts, what can be inferred about the customer's visitation behavior and what she may do in the current and future visits. Our modeling approach also differs from the above papers in that we assume that a customer develops a latent "inventory of experience" as she goes through her path of contacts with the retailer. The observable activity of a customer is stochastically dependent on this latent inventory. Given a customer's current inventory, we model her probabilities of purchase and continuation of browsing. In the case of continued browsing, we model when the customer is expected to return and through which channel.

Another key difference between our research and previous work is that we develop a model that is amenable to application to data that a typical, moderately sophisticated seller can collect very easily. Specifically, we *only* use data from prospects of the retailer who *visit* the website. Conditional on a visit, we know which channel the customer came from (e.g., by clicking on a link on a search engine, by clicking on an email, by directly typing the name of retailer in the browser's address bar, etc.) and whether she purchased or not in that visit. The other papers mentioned use richer datasets that, in turn, are also more difficult to collect. Specifically, the other papers have datasets obtained by tracking a panel of specific users over time, including tracking some of their activity outside the focal website (for instance, tracking exposure to display ads that were not clicked). Such data, though very useful in helping to answer path to purchase questions, need a significant amount of effort to collect. Indeed, these data can be difficult to collect even for fairly sophisticated firms. For instance, our conversations with a leading seller of books in the US market revealed that this firm, even though it has a large dedicated analytics team, finds it extremely difficult to collect data on exposures that did not lead to website visits. On the contrary, the type of data that we use in this paper (i.e., customer data conditional on a website visit), can be collected with almost no extra effort over what is expended by their analytics team in the usual order of business. Furthermore, our model can work with sparse navigation histories. Indeed, the data that we use for estimating the model in this paper is for a durable good for which browsing and purchase are infrequent because customers have relatively short records of visitation and conversion. This is in contrast to other papers mentioned that have richer data.

The rest of the paper is organized as follows. In the following section, we describe the methodology we propose to study the problem. In Section 3, we describe the data we use to demonstrate our approach in an empirical application. In Section 4, we present parameter estimates and interpret the results. In Section 5, we deconstruct channel contributions using analysis and simulations to describe how channels could play different roles in assisting conversions. In Section 6, we conclude with a discussion.

2. Modeling Approach

Our general modeling approach is based on the idea that observed behavior is a probabilistic outcome of a latent propensity to conduct that behavior. The model we build falls in the class of probability models of customer behavior (see Fader and Hardie (2009) for a review). The basic premise in the model is that each contact that the costumer has with the retailer affects *both* the customer's purchase probability *and* future browsing pattern, and this impact can depend on the whole history of contacts of the customer. In accordance with this, we model the impact of contacts through different channels on two types of latent propensities: the propensity to purchase given a contact and propensity to contact again (through the same or a different channel).

In particular, we define two "inventories of experience" that summarize the latent state of the customer at any point of time: $V_i(t)$ that modifies purchasing behavior and $W_i(t)$ that modifies browsing behavior. The observed behavior of a customer is a stochastic function of these latent inventory values. (For a similar modeling framework, see Moe and Fader (2004) and Jerath et al. (2015)). The values of these latent inventory variables depend on the number of times that each customer i has visited the website through channel k in each instant t, $N_{ik}(t)$. Thus, every time a customer uses a channel, we adjust the corresponding latent variable.

$$V_i(t) = \sum_k \rho_k^V \ln\left(1 + N_{ik}(t)\right) \tag{1}$$

$$W_i(t) = \sum_k \rho_k^W \ln\left(1 + N_{ik}(t)\right) \tag{2}$$

In the above, $\rho_k{}^V$ and $\rho_k{}^W$ are parameters that measure the impact of website visits through channel k on purchasing and browsing, respectively. The importance of keeping two separate accounts $V_i(t)$ and $W_i(t)$ is justified by the differential effect we expect them to play in each behavior. (We note that, in general, the inventory is a multidimensional construct depending on the different types of behavior that need to be modeled; in this application, we assume it to have two dimensions.) After each visit to the website, the two constructs are updated. The proposed specification allows for each channel to have a different impact in how consumers gather information and how they make their purchase decisions. The logarithm form captures the assumption that each additional visit to the site contributes less than the previous one. (The model is easily extendable to capture other shapes, such as a sigmoid shape, by simply using the appropriate functional form.)

First, we describe how we model the propensity of the customer to purchase from the retailer conditional on a visit. Let $p_i(t)$ be the probability that customer i buys from the retailer in the visit to the website at time t. (Given we do not observe competitive information we make no distinction between buying from someone else and simply not buying.) We describe this probability using a simple logit form as is described in equation (2).

$$p_i(t) = \frac{\exp\left(\rho_0^V + V_i(t)\right)}{1 + \exp\left(\rho_0^V + V_i(t)\right)} \tag{2}$$

We assume that the purchase decision is made at the end of a session, which is a cluster of one or more visits (details provided shortly). Consistent with the durable nature of the product being sold, in our data we observe that customers make no more than one purchase in the observation period. We therefore assume that after making a purchase, a customer leaves the market. However, extending our model to repeat purchases is straightforward.

Second, we describe how we model the propensity of the customer to use each of the channels. A customer could arrive at the website through one of multiple sources. For example, she could find the retailer using a search engine, she could enter the URL directly in the address bar, or she could respond to an email campaign. Patterns of online visits are typically characterized by *sessions* with multiple contacts with relatively short intra-visit times within a session and longer times between sessions (Park and Park 2016). We therefore consider that the intensity of the arrival process within a session is different from the intensity between sessions. The literature reports many alternative mechanisms to identify sessions in a sequence of customer visits to a website (He et al. 2002; Spiliopoulou et al. 2003; Huang et al. 2004). In our model, we use a simple, widely-used approach, under which two visits belong to the same session if the inter-arrival time is lesser than a session length τ (Jones and Klinker 2008). Creating sessions helps not only to accommodate lumpy patterns in visits, but also enables us to identify two different effects that channels have, namely on the inter-arrival and intra-arrival processes. While inter-arrival times are related to the ability of channels in generating new sessions, intra-arrival times address their abilities in extending sessions. We describe inter- and intra-session arrivals to the website from each channel by simultaneous Poisson processes. Here, the arrival rates for each customer i at time t are characterized by two vectors of rates: $\boldsymbol{\lambda}_i(t)=(\boldsymbol{\lambda}_{i1}(t),\ldots, \boldsymbol{\lambda}_{iK}(t))$ for the intra-session arrivals, and $\mu_i(t)=(\mu_{i1}(t),\ldots, \mu_{iK}(t))$ for the inter-session arrivals, where K is the number of channels available for the customers. Formally, the distribution of intra- and inter-session arrivals in a period of length d at a time t are given in equations (2) and (3) Should we Define Xik and Yik and d.

$$\Pr\left(X_{ik}(t,t+d)=x\right)=\left(\lambda_{ik}(t)d\right)^{x} e^{-\lambda_{ik}(t)d} \Big/ x! \tag{2}$$

$$\Pr\left(Y_{ik}(t,t+d)=y\right)=\left(\mu_{ik}(t)d\right)^{y} e^{-\mu_{ik}(t)d} \Big/ y! \tag{2}$$

Figure 1 illustrates how our proposed model describes visits to the website. When an inter-session arrival happens at time t, the intra-session process starts. If a new visit occurs before $t+\tau$, then this session continues; accordingly, we update the arrival rates and a new intra-session arrival process starts. If no visit occurs before $t+\tau$, then a new inter-session arrival process starts at $t+\tau$. Our description of customer visitations corresponds to a class of competing hazard rate models. Other recent marketing applications use similar approaches (Braun and Schweidel 2011; Moe and Trusov 2011). Can we match this explanation with Figure 1?

(Insert Figure 1 about here)

It is important to note that both intra- and inter-session arrival rates depend on the latent inventory of experience $W_i(t)$ which is modified every time the customer visits the website. Each channel is characterized by its own propensity and those propensities can be modified as a function of the browsing history. For example, after using a search engine to search for a brand name, the propensity to enter the URL directly in the address bar could increase, while the propensity to conduct a search at a search engine could decrease.

Next, we specify the parametric form of the arrival rates for channel k, for customer i for whom the last visit occurred using channel h. For simplicity, we adopt an additive model with three components: a base rate $\rho_0{}^W$ that captures the intrinsic propensity to visit the website, the inventory of experience $W_i(t)$ to capture the cumulative effect of browsing history, and ω_{hk} to capture the sequence of channels that the customer uses to visit the website. We expect to observe different patterns of navigation between and within sessions and therefore we have different parameters for each case. Can we explain the connect between the Mu in equation (4) and the Mu in equation (3)? I got lost here.

$$\mu_{ihk}(t) = \exp\left(\rho_0^W + W_i(t) + \omega_{hk}^{\mu}\right) \tag{2}$$

$$\lambda_{ihk}(t) = \exp\left(\rho_0^W + W_i(t) + \omega_{hk}^{\lambda}\right) \tag{2}$$

There are multiple alternative models to describe browsing sequences. For simplicity, we focus on first order Markovian models where, given the cumulative stocks of inventory, the likelihood of using each channel depends only on the last channel used. Extensions to cases where longer sequences are explicitly incorporated into the model are conceptually straightforward but dramatically increase the number of parameters to be estimated. We propose an additive form with parameters depending on the last channel visited and the potential channels to be visited:

$$\omega_{hk}^{\mu} = \alpha_h + \beta_k + c_k \delta_{hk} \tag{2}$$

$$\omega_{hk}^{\lambda} = \kappa_h + \eta_k + d_k \delta_{hk} \tag{2}$$

where δ_{hk}=1 if h=k.

We interpret parameters β and η as representing channel *attractiveness*: this is the drawing power that each channel has regardless of the browsing history. For example, customers of a retailer with little brand awareness are not likely to include the brand name in search engine queries and therefore generic keywords constitute a natural search choice. In this case, we would expect larger attractiveness parameters for competitive search than for brand search. Similarly, we interpret parameters α and κ as representing channel *engagement*, or the ability of the channel to keep the customer interested to generate subsequent visits in the future. For example, when customers conduct price comparisons, they are likely to obtain information about direct competitors, which can divert customer attention thus decreasing the likelihood of continuing their purchase process within the focal retailer. On the other hand there is no presence of competitive information in channels such as emails or direct load, for which we expect larger engagement parameters. For identification purposes, we impose β_K=η_K=0 (i.e., we fix the values of these parameters for one channel, which is the channel identified by the label K).

Finally, parameters c and d are included to capture channel *persistence*, or the tendency to use the same channel that was used in the most recent visit (Montgomery et al. 2004). We expect that c, d>0 implying that after using channel k, the likelihood of using the same channel in the next visit increases. Compared to the full model with one parameter per pair of channels to describe transition probabilities, which would require $2K(K\text{-}1)$ parameters, this is a parsimonious specification with $6K\text{-}2$ parameters.

To derive the likelihood of the model, it is useful to recognize that the state variables $W_i(t)$ and $V_i(t)$ only change when the customer visits the website. Therefore, we can decompose the individual likelihood per period ($t=\{1,2,\ldots, T_i\}$). We define each period as the time between two contacts. Moreover, we recognize that contacts defining a period can be classified into four categories with relatively simple likelihood contributions. We use these components, together with the purchase model, to write the full likelihood. We define the following auxiliary variables to write the likelihood. Let x_{ikt} be a dummy variable taking the value 1 if customer i uses channel k for her visit to the website in period t (0 otherwise). Similarly, let y_{it} be an indicator variable taking the value 1 if the customer makes a purchase in her visit to the website in period t. To complete the notation, let d_{it} be the duration of the period t for the customer i.

(I) FIRST ARRIVAL: We assume that the first time the customer arrives to the website, she initiates a new session and therefore we consider inter-session arrival rates only. By definition, in each period we observe exactly one occurrence for one channel and none for all others. The log-likelihood of the channels with no visit is just the effective rate of the Poisson process given by the product of the arrival rate and the duration of the period. For the channel with an occurrence we need to add the logarithm of the effective rate.

$$LL_{it}^{(1)} = \ln\left(\sum_k x_{ikt}\mu_{ik}(1)d_{it}\right) - \sum_k \mu_{ik}(1)d_{it} \tag{2}$$

(II) INTERIOR INTER-SESSION ARRIVAL: Unlike the previous case, we need to decompose the inter-visit time into two parts. First, an interval of length τ where no intra-session visits occur and, second, an interval of length $(d_{it}\text{–}\tau)$ where only one inter-session visit occurs.

$$LL_{it}^{(\mathrm{II})} = \ln\left(\sum_k x_{ikt}\mu_{ik}(t_i)\left(d_{it}-\tau\right)\right) - \sum_k \lambda_{ik}(t_i)\tau - \sum_k \mu_{ik}(t_i)\left(d_{it}-\tau\right) \tag{2}$$

(III) INTERIOR INTRA-SESSION ARRIVAL: This case follows the same logic as for the first visit, but considering intra-session arrival rates.

$$LL_{it}^{(\mathrm{III})} = \ln\left(\sum_k x_{ikt}\lambda_{ik}(t_i)d_{it}\right) - \sum_k \lambda_{ik}(t_i)d_{it} \tag{2}$$

(IV) AFTER LAST ARRIVAL: This case is similar to the interior inter-session arrival case, but with all channels having no visits. Additionally, we need to truncate the duration in case it is longer than the remainder of the observation period. Given that we deal with durable goods, we assume that customers who made a purchase are out of the market and therefore they are not considered in this expression.

$$LL_{it}^{(\mathrm{IV})} = \left(1-y_{iT_i}\right)\left[-\sum_k \lambda_{ik}(T_i)\min\{\tau, d_{iT}\} - \sum_k \mu_{ik}(T_i)\max\{d_{iT}-\tau, 0\}\right] \tag{2}$$

(V) PURCHASE: To complete the likelihood we need to write the contribution to the likelihood of purchase behavior.

$$LL_{it}^{(\mathrm{V})} = y_{it}\left[\left(\rho_0^V + V_i(t)\right) - \ln\left(1+\exp\left(\rho_0^V + V_i(t)\right)\right)\right] \tag{2}$$

Then, the full likelihood of the model is given by

$$LL = \sum_{it\in \mathrm{I}} LL_{it}^{(\mathrm{I})} + \sum_{it\in \mathrm{II}} LL_{it}^{(\mathrm{II})} + \sum_{it\in \mathrm{III}} LL_{it}^{(\mathrm{III})} + \sum_{it\in \mathrm{IV}} LL_{it}^{(\mathrm{IV})} + \sum_{it\in \mathrm{V}} LL_{it}^{(\mathrm{V})} \quad (2)$$

3. Data and Estimation

Our data are obtained from an online retailer that tracks customers who visit its website. For every visit, the marketing channel that the visit originated from is recorded, and whether the customer made a purchase in this visit is also recorded. The data set corresponds to a sample of 20,000 customers visiting the website in a single shopping season (15 weeks). The firm sells durable goods and at most one purchase is made per customer in this observation period. The list of sources from which the contacts originate is provided in Table 1, along with the percentage of customer contacts made using each of them.

(Insert Table 1 about here)

In our data, the only communication vehicle that is directly initiated by the company is email contacts. An obvious limitation of the data is that we only observe those customers who visit the website through at least one of the touch points. For example, if a customer received an email, but did not open the email, then we would not have any observations in our data set. On the other hand, these data are extremely easy for any firm to collect while doing "business as usual," i.e., no significant special arrangements have to be made to collect these data, which increases the relevance and applicability of our model.

Many of the customers in our dataset visit the website only once in the observation time period, and only a small percentage of them actually purchase from the retailer. However, a significant fraction of customers have multiple contacts using multiple channels; these data help us to extract valuable information about online shopping behavior. We note that customers who end up purchasing on the website have 3.44 visits on average, whereas non-buying visitors only have 1.62 visits on average.

To define session lengths we set τ=8hrs. We conduct sensitivity analysis over session lengths with values of τ=2, 4 and 12 hours and, consistent with Montgomery and Faloutsos (2007), we find that all inferences remain unchanged.

4. Results

4.1 Model Evaluation

We evaluate the performance of the model using multiple metrics. We first compare the model against three simpler specifications to establish that not only is tracking visitation history important, but also that grouping into sessions significantly helps to describe browsing and purchase behavior. First, we consider a model where each channel has constant intra- and inter-session arrival rates and a constant per-session purchase probability. Compared to this model, ours has better performance in terms of goodness of fit (AIC=321,187.3 vs. AIC= 369,491.3). We then compare our model against a nested model where we impose that the history of visitation has no impact on customer behavior. This model has even worse fit (AIC=369,543.3) and, using the likelihood ratio test, we reject the null hypothesis $\rho^V=\rho^W=0$ (p-value<0.001). Finally, we compare against a model in which there is no distinction between inter- and intra-session arrival rates. This final benchmark also perform much worse (AIC=403,281.6) and consequently, likelihood ratio test rejects the null that $\lambda_{hik} = \mu_{hik}$ (p-value<0.001). Based on these comparisons we conclude that separating into sessions, allowing different inter- and intra-session rates, and allowing these rates to be dependent on visitation history are all components that improve the model's in-sample fit.

To further evaluate how the model describes visitation behavior, we analyze its ability to replicate aggregate visit numbers for each channel. More specifically, we build histograms of the number of visits per channel and compare them against the actual numbers, as shown in Figure 2. For each channel, we cannot reject the null that the observed data comes from the same distribution than the one governed by parameter estimates (all p-value>0.9 for Pearson's χ^2 test) providing additional support for the model.

(Insert Figure 2 about here)

Although the model fits reasonably well and the means are all correctly captured, it is worth noting that there is room to accommodate more complex models with more flexible arrival processes. For example, the distribution for Direct Load from our model has a heavier tail than we observe in the data. On the other hand, the distribution for Search Competitive has a lighter tail than observed in the data.

Finally, we evaluate the ability of the model to predict customer behavior on a holdout sample. After each contact, a customer could engage in 15 different behaviors: extending the current session using any of the 7 available channels, not extending and buying, and finally not extending nor buying but initiating a new session in the future using any of the 7 available channels. For each observed contact in the sample, we compute the probability of occurrence for each of these actions and we then compare them against the proportion with which they actually happen. In this additional exercise we use the first 84 days of data to calibrate the model and the reminding 21 to validate. As shown in Figure 3, the model seems to do a reasonable job in replicating aggregated behavior. The mean absolute percentage error (MAPE) in the validation sample at this level of aggregation is 11.1%[3]

(Insert Figure 3 about here)

4.2 Parameter Estimates

We present parameter estimates in two parts. We start by describing the contribution that each channel has in the latent inventories of experience. The maximum likelihood estimates and the corresponding standard errors are reported in Table 2.

[3] We note that visually the model does a good job; however, the comparison of predicted probabilities and actual frequencies does not pass a Pearson's χ^2 test (p-value<0.001). This can be explained by two reasons. First, given the large number of observations in the dataset even small relative deviations from the expected values generate large χ^2 statistics (if we simply sample 1,000 observations from the data then we cannot reject the null that both series comes from the same distribution, p-value > 0.01). The second reason is given by the structure we impose in the model where intra- and inter-arrival rates are not independent. For example, the channel with the largest error is Direct Load where session extensions are underestimated, but new session initiations are overestimated. A less parsimonious model with complete independence between these processes could increase intra-arrival rates and decrease inter-arrival rates for Direct Load.

(Insert Table 2 about here)

On examining Table 2, we can see that the baseline probabilities of purchase and arrival are quite small (because the corresponding estimates are large and highly significant negative numbers). This is consistent with a durable product that customers shop for and purchase infrequently. For the other estimates, a positive number indicates that the corresponding channel leads to a probability greater than baseline, and a negative number indicates that the corresponding channel leads to a probability smaller than baseline. Examining the contributions that channels make in generating a sale (parameters with superscript V), we find that Search Brand, Other Paid and Direct Load have the larger values. The effects that these channels have in terms of instantaneous conversions lead us to characterize them as playing the role of "closers." In terms of affecting visitation (parameters with superscript W), we find moderate differences across channels, with Search Competitive, Search Brand and Shop Engine being the channels that decrease the intensity of visits the most. It is worth noting that these channels are precisely the ones that provide competitive information to the customer. These results indicate that the value proposition of this retailer may be somewhat weak in the face of comparisons with competitors.

(Insert Tables 3a and 3b about here)

Tables 3a and 3b display the parameters associated with browsing behavior for inter- and intra-session browsing, respectively. We first consider channel attractiveness; recall that β and η are inter- and intra-channel attractiveness parameters, respectively. Comparing β values in Table 3a, we find that Search Competitive is the most attractive channel to initiate a new session. However, comparing η values in Table 3b, we find that within a session the most attractive channel is Direct Load. These results suggest we are dealing with a retailer with little brand awareness that attracts customers at the beginning of their purchase processes mainly through competitive search.

Once the retailer's brand is in the customer's mind, they start using it by typing it directly in the address bar.

We first consider channel engagement; recall that α and κ are inter- and intra-channel engagement parameters, respectively. Comparing κ values in Table 3b, we observe that there is little difference across channels within a session. The only one with slightly smaller engagement is Email, which can be explained by a lack of coordination between the time the email campaign is sent and the time the customer actually wants to conduct online research about the product. Comparing α values in Table 3a, we observe that there is more variation across channel effects between sessions. Somewhat surprisingly, Search Competitive has the largest likelihood of generating a new visit in the future. The channels that have larger effects in terms of motivating future browsing lead us to characterize them as playing the role of "engagers." It is interesting to note that Email has the lowest likelihood of keeping the customer navigating in the current session but does not have the lowest likelihood of starting a new session in the future when the customer is really interested in initiating an active search process.

There are other interesting aspects of browsing behavior that we can identify from our results. Recall that persistence, or the tendency to use the same channel as was used in the most recent visit, is captured by parameters c and d for the inter- and intra-session cases, respectively. Given that the estimates of c and d are mostly significant and positive, we conclude that channel persistence is high. Moreover, our results show that, in general, $c<d$, implying larger persistence within sessions than between sessions.

To further understand browsing dynamics, it is useful to compute transition matrices that describe the probability of using each channel conditional on the identity of the current one being used. Given that we describe the arrival processes using a Poisson processes, the probability that the next visit comes from a specific channel is simply the ratio of that channel arrival rate to the sum of all rates (Ross 2003). We report this in a channel switching probability matrix describing the likelihood of using each channel after using any other channel. Since we have different parameters for arrivals within and between sessions, we report separate switching probability matrices for intra- and inter-session arrivals, given in Tables 4a and 4b, respectively.

(Insert Tables 4a and 4b about here)

In Tables 4a and 4b, large numbers in the diagonal imply that there is high persistence in channel usage, i.e., customers tend to use the same channel they used in previous contacts. We also find that using the same channel in two consecutive visits is more likely within a session than between a session. This is justified by shorter inter-visit times where inertia in channel choice is more prominent (Jeuland 1979; Roy et al. 1996). There are other interesting patterns as well in the sequences of channel choices. For example, a direct load of the company URL in the address bar is a very common choice after using any of the other channels. However, to start a new session costumers prefer using Search Competitive. This is consistent with the description of a retailer with relatively low brand awareness.

5. Channel Role and Contribution

Our results above provide a number of interesting insights on the relative roles of the different channels through which consumers visit. In addition, the real power and usefulness of our model lies in its ability to generate various forward-looking metrics to compare the efficacy of different channels. Our model enables us to identify at least three important effects that channels might have in modifying consumer behavior—the ability to generate new sessions, the ability to extend sessions and finally the instantaneous effect they could exert in generating a sale. To understand the joint impact of these components, we need to integrate them by building simple mechanisms that managers can use to make decisions regarding multichannel marketing strategies. The key challenge for doing so is generating a metric of purchase probability accounting for the future purchase process, and not just the current visit.

Using our model, we can compute the probability the customer will make a purchase at any point of time in the horizon. We provide two metrics of channel contribution. First, we consider the *short-term* purchase probability, which corresponds to the probability of buying in the current or in the next visit to the website.

Second, we consider the *medium-term* purchase probability, which corresponds to the probability of purchasing during the full shopping season under consideration, i.e., fifteen weeks.[4] It is important to note that the arrival rates and purchase probabilities depend on previous contacts and, therefore, the marginal impact of a customer visitation depends on the browsing history. Thus, we need to average the impact that a particular visit has over different paths. More precisely, to get this average effect, we need to account for multiple possible trajectories of visitations of customers as represented in the dataset. This gives the empirical endowments of inventory observed in the dataset. Specifically, we take a random sample of customers from the database to simulate from representative initial endowments of experience. For a given customer, given her current endowment, we use the model to generate one exponentially distributed random variable for each channel representing the next arrival time and we take the minimum of these as the realization of the next visit. We update the endowment and the purchase probability and then repeat until reaching the simulation horizon. Using this simulation procedure, we compute a baseline value for the comparison. Next, to assess the marginal impact of one additional visit from a specific channel for this customer, we simulate again in the same way as above, but modify the initial

[4] We do not consider a long-term impact because this would go beyond the shopping season. To make reasonable calibrations for the long term, we would need data for a longer time period.

endowment by a value equivalent to one additional impact of the channel under consideration. Taking the difference between the purchase probabilities of this second simulation with respect to the baseline, we obtain the medium-term impact of one additional contact through this channel for this customer. We repeat this process for the random sample of customers, which gives us the medium-term impact of one additional contact through this channel for all customers in the sample. Finally, we can obtain the impact of this channel for the customer base by averaging the impact across all customers.

Table 5 displays short- and medium-term impact derived from the model. Negative numbers in the two columns are interpreted to mean that a customer using the corresponding channel is less likely to purchase compared to a similar customer who does not use the channel but is the same in every other way. In general, the model suggests that, for our application, channels have a negative or relatively small impact in generating future sales. This could be explained by the fact that this is a durable good purchased typically once every season. In addition, it could also indicate that the competitive positioning of the retailer is weak, and the retailer is not very effective in keeping customers engaged. Nevertheless, for some channels the effect is not negligible. For example, if a customer makes an additional visit to the site through the channel Other Paid, her likelihood of making a purchase in the current or next visit increases on average by 0.007 which corresponds to a 11.8% lift from the average purchase probability of 0.059.

(Insert Table 5 about here)

Figure 4 displays a bubble chart with the short-term impact on the horizontal axis and the midterm-term impact on the vertical axis. The size of each bubble represents the fraction of visits made through that channel. We note that managing media by trading off volume versus effectiveness in both the short term and the long term is a classic marketing problem, and Figure

4 provides an intuitive way of summarizing this. As expected, there is a positive correlation between short- and medium-term impacts. However, we do observe some reversals between short- and medium-term evaluations. For example, Other Paid has a large impact on generating immediate sales, but if the sale is not made now, it implies a relatively low impact in the medium term. This is consistent with the role of “closer” we assigned to this channel before. In spite of the reversals, we observe some general consistency in the channel ranks as a function of the time frame. Among channels having consistently the largest impact in generating sales are Direct Load and Search Brand. Interestingly, both channels denote some level of brand awareness from the customer. Thus, the use of these channels could signal customer prior interest in buying or an intrinsic preference for the retailer.

(Insert Figure 4 about here)

Search Competitive and Shop Engine have small impact on sales for both short and medium term. Interestingly, both channels are the ones providing information regarding the competitors to the customer. Another interesting case is that of the email channel that has one of the largest negative effects in the short term, but in the medium term this reverses to a mild positive effect. This is consistent with the nature of the email communications that are primarily firm initiated. Therefore, for the category analyzed, Email does not help to generate immediate sales, but when considering a longer time horizon, Email has a relatively larger likelihood of starting a new session in the future. Other Free also ranks low on both short and medium term in generating sales, which could also be associated with uncoordinated timing.

Next, we decompose the impacts of channels into extending sessions, returning to initiate a new session, and purchase. Using the simulation explained before, we compute the above components and in Figure 5 display the boxplots of the marginal impacts in extending sessions, returning to initiate a new session and purchase for each of the customers in the sample. For example, a value of 0.2 in the first column of panel (a) means that for that customer an additional endowment of a Direct Load channel generates a positive increase of 0.2 in the probability of extending the session. Note that the different points for a given channel correspond to the simulated values for different individuals with different endowments. The variation shows that the impact can vary substantially with endowment.

(Insert Figure 5 about here)

This decomposition confirms several of the implications of the previous analyses related to the different roles played by different channels. For example, for the Email channel, it is interesting to note that even though an additional contact using the Email channel decreases the probability of extending a session the most, it is one of the channels with the best relative performance in initiating new sessions. A narrative that is consistent with these results is that emails, because they are firm initiated, may be received when customers are not ready to initiate an active purchase process, but provide enough awareness to return to the website in the future. Direct Load has almost no effect in extending sessions, but once a customer types the website URL in the address bar, it indicates that the customer has started a new session with the retailer. In terms of immediate purchase probability, we observe that the two channels having the largest impact are Other Paid and Search Brand. In our application, it can be explained by the relatively short browsing sequences.

6. Conclusions and Discussion

Understanding the interconnected roles played by different channels in impacting consumers in a multichannel setting is a practical marketing problem that firms face in today's environment. In this paper, we develop a probability modeling-based approach as a solution to this problem. Our model accounts for a customer's full browsing history and summarizes it into an "inventory of experience." This inventory is a multidimensional latent construct, and observed behavior is a stochastic function of the latent construct. We model purchasing and browsing behavior of a customer. Furthermore, for browsing behavior, we model intra-session and inter-session browsing while accounting for transitions across channels. Taken together, these components help to obtain estimates of the relative values of each channel. Our model allows us to estimate the short-term and medium-term impact of each channel. We develop our method such that it can be applied on a data structure that any firm with a website and basic analytics functionalities would find easy to collect.

We estimate our model using data from a pure-play online retailer selling a durable good. We find that different channels play different roles and can be characterized into "closers" and "engagers." Customers arriving through channels characterized as "closers" (which include Direct Load and Search Brand) are more likely to make a purchase in the current visit, while customers arriving through channels characterized as "engagers" (which include Email and Search Competitive) are more likely to visit again in the future, even if they don't make a purchase in the current visit. Therefore, while some channels have value because they are strong on generating instantaneous conversions, others have value because they increase the likelihood of generating new visits to the website in the future. We also find that the browsing history of a customer has an important role to play and ignoring this history can generate incomplete views on customer behavior and incorrectly estimate the channel order in terms of the channels' abilities to generate value to the company. We also find very strong persistence across visits in which channel a customer arrives from; this might indicate that customers have strong intrinsic channel preferences.

We conclude by identifying some promising extensions of our research for future investigation. First, the literature on empirical analysis of dynamic decisions frequently considers that more recent events have a larger effect in modifying current behavior (Lattin and Bucklin 1989, Mehta et al. 2004). We expect that incorporating forgetting mechanisms in our application could enrich our description of the role of multiple channels in customer behavior. Second, in the base model we concentrate on whether the customer buys or not. An interesting extension of this would be to include the expenditure decision or to model other micro conversions such as signing up for email or ordering samples. Third, as pointed out by Schmittlein et al. (1987), a customer might become inactive during the period of her relationship with a company, and a proper description of the process by which the customer becomes inactive can help to describe her behavior. The literature posits many alternative models to describe the process (e.g., Jerath et al. 2011). The relatively short period for which we have data, accompanied by the fact that we model purchase of a durable product, makes the dropout process somewhat unnecessary to model. However, the application of our model to other settings might require this incorporating a dropout component. Finally, we essentially model correlations between channel usage and conversions in this paper. But do these correlational patterns indicate intrinsic channel preferences of customers, or do customers change their behavior after using a channel. In support of the former argument, the fact that the Shopping Engine channel has a negative effect could be explained by positing that more price sensitive customers using that channel. In support of the latter argument, we note that, in our application, all channels that provide competitive information generate similar browsing and purchase patterns. An interesting extension of this work would be to disentangle these two explanations by either adding more structure to the model, or using richer data, or by using experimental methods where channel activity is exogenously varied.

References

Abhishek, Vibhanshu, Peter S. Fader, and Kartik Hosanagar (2013) “Media Exposure through the Funnel: A Model of Multi-Stage Attribution,” Working paper, Carnegie Mellon University, Pittsburgh.

Anderl, Eva, Ingo Becker, Florian v. Wangenheim, and Jan H. Schumann (2013), “Putting Attribution to Work: A Graph-Based Framework for Attribution Modeling in Managerial Practice,” Working paper, Passau University.

Braun, Michael, and David A. Schweidel (2011), "Modeling Customer Lifetimes with Multiple Causes of Churn." *Marketing Science*, 30 (5), 881-902.

Fader, Peter S., and Bruce GS Hardie (2009), “Probability Models for Customer-Base Analysis.” *Journal of Interactive Marketing*, 23 (1), 61–69.

Fader, Peter S., Bruce G. S. Hardie and Ka Lok Lee (2005), “RFM and CLV: Using Iso-Value Curves for Customer Base Analysis,” *Journal of Marketing Research*, 42 (4), 415–430.

He, Daqing, Ayşe Göker, and David J. Harper (2002), “Combining Evidence for Automatic Web Session Identification,” *Information Processing & Management*, 38 (5), 727–742.

Huang, Xiangji, Fuchun Peng, Aijun An, and Dale Schuurmans (2004), “Dynamic Web Log Session Identification with Statistical Language Models,” *Journal of the American Society for Information Science and Technology*, 55 (14), 1290–1303.

Jeuland, Abel P (1979), “Brand Choice Inertia as One Aspect of the Notion of Brand Loyalty,” *Management Science*, 25 (7), 671–682.

Jerath, Kinshuk, Peter S. Fader, and Bruce GS Hardie (2010), "New Perspectives on Customer 'Death' Using a Generalization of the Pareto/NBD Model," *Marketing Science*, 30 (5), 866–880.

____, Anuj Kumar, and Serguei Netessine (2015), "An Information Stock Model of Customer Behavior in Multichannel Customer Support Services," *Manufacturing & Service Operations Management*, 17 (3), 368–383.

Jones, Rosie, and Kristina Lisa Klinkner (2008), "Beyond the Session Timeout: Automatic Hierarchical Segmentation of Search Topics in Query Logs," in *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, 699–708.

Lattin, James M., and Randolph E. Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior," *Journal of Marketing Research*, 26 (3), 299–310.

Laukkanen, Tommi (2007), "Internet vs Mobile Banking: Comparing Customer Value Perceptions." *Business Process Management Journal*, 13 (6), 788–797.

Li, Hongshuang, and P. K. Kannan (2014), "Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment," *Journal of Marketing Research*, 51 (1), 40–56.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2004) "Role of Forgetting in Memory-Based Choice Decisions: A Structural Model," *Quantitative Marketing and Economics*, 2 (2), 107–140.

Moe, Wendy W., and Peter S. Fader (2004), "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science*, 50 (3), 326–335.

____, and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48 (3), 444–456.

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), 579–595.

Montgomery, Alan L., and Christos Faloutsos (2007), "Identifying Web Browsing Trends and Patterns," *IEEE Computer*, 34 (7), 94–95.

Park, Chang Hee, and Young-Hoon Park (2016), "Investigating Purchase Conversion by Uncovering Online Visit Patterns," *Marketing Science*, 35 (6), 894–914.

Payne, Adrian, and Pennie Frow (2005), "A Strategic Framework for Customer Relationship Management," *Journal of Marketing*, 69 (4), 167–176.

Ross, Sheldon M. (2003) *Introduction to Probability Models, Eighth Edition*. Academic Press.

Roy, Rishin, Pradeep K. Chintagunta, and Sudeep Haldar (1996), "A Framework for Investigating Habits, "The Hand of the Past," and Heterogeneity in Dynamic Brand Choice." *Marketing Science,* 15 (3), 280–299.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987). "Counting Your Customers: Who Are They and What Will They Do Next?", *Management Science*, 33 (1), 1–24.

Spiliopoulou, Myra, Bamshad Mobasher, Bettina Berendt, and Miki Nakagawa (2003), "A Framework for the Evaluation of Session Reconstruction Heuristics in Web-Usage Analysis," *INFORMS Journal on Computing*, 15(2), 171–190.

Wallace, David W., Joan L. Giese, and Jean L. Johnson. (2004). "Customer Retailer Loyalty in the Context of Multiple Channel Strategies," *Journal of Retailing*, 80 (4), 249–263.

Xu, Lizhen, Jason A. Duan, and Andrew Whinston (2014), "Path to Purchase: A Mutually Exciting Point Process Model for Online Advertising and Conversion," *Management Science*, 60 (6), 1392–1412.

Channel (Code)	Description	All	Buyers	Non-buyers
Direct Load (DL)	The customer visits the website by typing in the firm's address directly into the address bar or by using a bookmark.	28.59 %	45.84 %	27.96%
Search Competitive (SC)	The customer clicks on the website link on a search engine after a search related to the category but not containing the firm's brand name.	33.87 %	16.04 %	34.53%
Search Brand (SB)	The customer clicks on the website link on a search engine after a search containing the brand name of the firm.	7.50%	9.59%	7.43%
Shopping Engine (SE)	The customer clicks on the website link from a comparison shopping engine site (such as Shopping.com or BizRate.com).	13.21 %	8.40%	13.39%
Email (EM)	The customer visits the website by clicking on the website link in an email by the firm.	7.75%	7.47%	7.76%
Other Paid (OP)	The customer clicks on the website link on an affiliate website, banner or other media which is paid for by the firm.	4.58%	8.82%	4.43%
Other Free (OF)	The customer arrives from a source not included in the above and that the company has not paid for, e.g., the website link on a forum such as Craigslist.	4.48%	3.83%	4.50%

Table 1: Description of channels. (The last three columns show the percentage of clicks that are accounted for by each channel for all customers, for buyers only, and for non-buyers only.)

		m.l.e.	s.e.
ρ_0^V	Base Purchase	-4.214	0.064
ρ_k^V	Direct Load	0.422	0.050
	Other Paid	1.196	0.133
	Search Competitive	0.033	0.095
	Search Brand	1.072	0.115
	Shop Engine	0.234	0.112
	Email	-0.067	0.106
	Other Free	-0.409	0.218
ρ_0^W	Base Arrival	-14.685	0.031
ρ_k^W	Direct Load	0.003	0.013
	Other Paid	-0.531	0.055
	Search Competitive	-1.394	0.028
	Search Brand	-0.603	0.038
	Shop Engine	-1.006	0.037
	Email	0.175	0.017
	Other Free	0.185	0.030

Table 2: Estimates of "inventory of experience"

		m.l.e.	s.e.
	Direct Load	-0.508	0.066
	Other Paid	-0.727	0.190
	Search Comp.	0.262	0.073
α	Search Brand	-0.083	0.097
	Shop Engine	0.093	0.104
	Email	-0.844	0.079
	Other Free	-0.959	0.143
	Direct Load	1.621	0.034
	Other Paid	0.053	0.044
	Search Comp.	2.028	0.033
β	Search Brand	0.622	0.039
	Shop Engine	1.108	0.036
	Email	0.409	0.040
	Other Free	0	fixed
	Direct Load	2.248	0.075
	Other Paid	2.533	0.288
	Search Comp.	0.748	0.091
c	Search Brand	2.034	0.139
	Shop Engine	1.888	0.133
	Email	3.458	0.094
	Other Free	3.181	0.191

Table 3a: Inter-session arrival parameter estimates

		m.l.e.	s.e.
	Direct Load	4.500	0.097
	Other Paid	5.399	0.132
	Search Comp.	5.715	0.096
κ	Search Brand	5.139	0.108
	Shop Engine	5.470	0.108
	Email	3.231	0.119
	Other Free	3.907	0.138
	Direct Load	2.478	0.089
	Other Paid	0.282	0.118
	Search Comp.	1.875	0.094
η	Search Brand	0.252	0.121
	Shop Engine	1.004	0.102
	Email	0.016	0.128
	Other Free	0	fixed
	Direct Load	1.994	0.072
	Other Paid	3.752	0.138
	Search Comp.	2.617	0.072
d	Search Brand	3.340	0.122
	Shop Engine	3.356	0.095
	Email	4.550	0.139
	Other Free	4.370	0.141

Table 3b: Intra-session arrival parameter estimates

	Direct Load	Other Paid	Search Competitive	Search Brand	Shop Engine	Email	Other Free
None	0.240	0.050	0.360	0.088	0.143	0.071	0.047
Direct Load	0.749	0.017	0.119	0.029	0.047	0.024	0.016
Other Paid	0.152	0.398	0.228	0.056	0.091	0.045	0.030
Search Competitive	0.171	0.036	0.543	0.063	0.102	0.051	0.034
Search Brand	0.151	0.031	0.227	0.425	0.090	0.045	0.030
Shop Engine	0.133	0.028	0.200	0.049	0.525	0.040	0.026
Email	0.075	0.016	0.113	0.028	0.045	0.709	0.015
Other Free	0.114	0.024	0.172	0.042	0.069	0.034	0.545

Table 4a: Inter-session transition probability matrix

	Direct Load	Other Paid	Search Competitive	Search Brand	Shop Engine	Email	Other Free
Direct Load	0.863	0.013	0.064	0.013	0.027	0.010	0.010
Other Paid	0.147	0.698	0.081	0.016	0.034	0.013	0.012
Search Competitive	0.110	0.012	0.823	0.012	0.025	0.009	0.009
Search Brand	0.196	0.022	0.107	0.597	0.045	0.017	0.016
Shop Engine	0.118	0.013	0.064	0.013	0.772	0.010	0.010
Email	0.099	0.011	0.054	0.011	0.023	0.795	0.008
Other Free	0.115	0.013	0.063	0.012	0.026	0.010	0.761

Table 4b: Intra-session transition probability matrix

	Short- Term Impact	Medium- Term Impact
Direct Load	0.000	0.005
Other Paid	0.007	-0.005
Search Competitive	-0.006	-0.009
Search Brand	0.002	0.000
Shop Engine	-0.003	-0.006
Email	-0.013	-0.004
Other Free	-0.014	-0.009

Table 5: Channel impact metrics. (The shaded columns report metrics obtained from other models, and the non-shaded columns report metrics obtained from our model.)

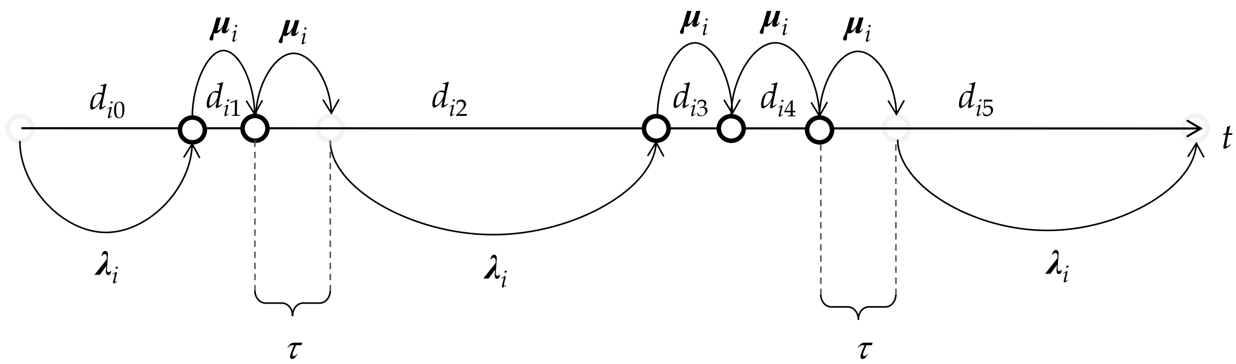

Figure 1: Illustration of customer visitation model

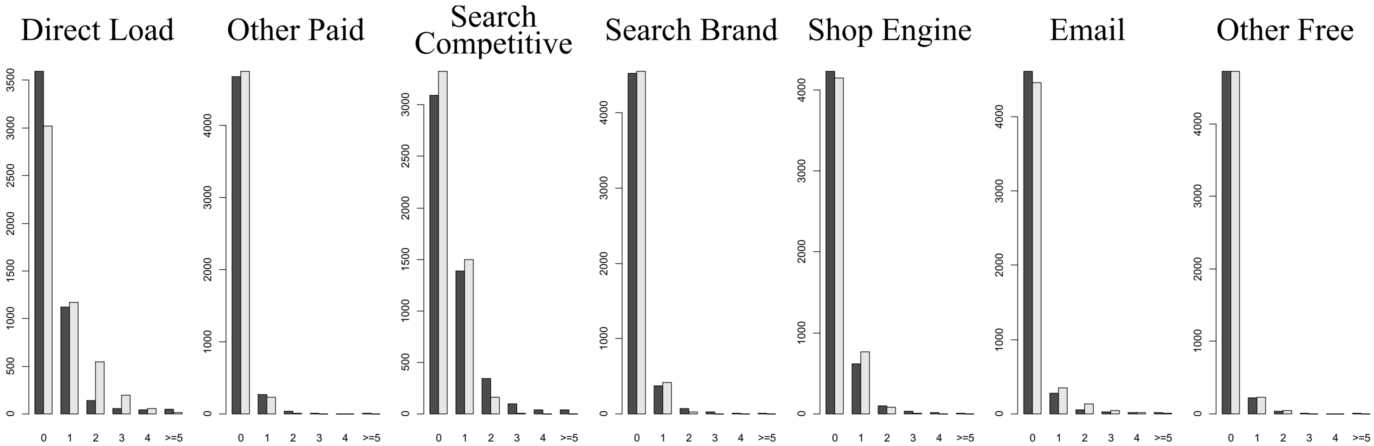

Figure 2: Actual (dark) vs. predicted (light) numbers of visits per channel.

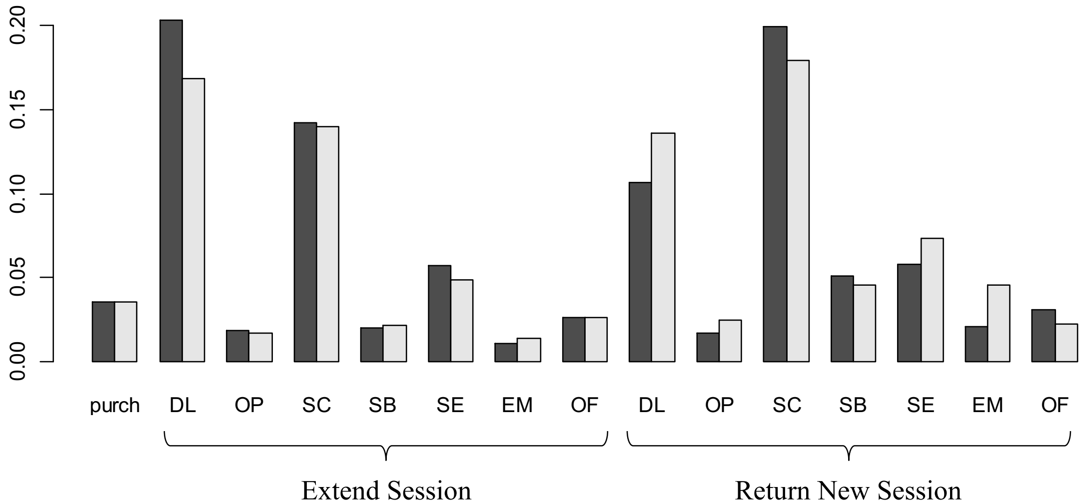

Figure 3: Actual (dark) vs. predicted (light) probabilities of customers' next behavior.

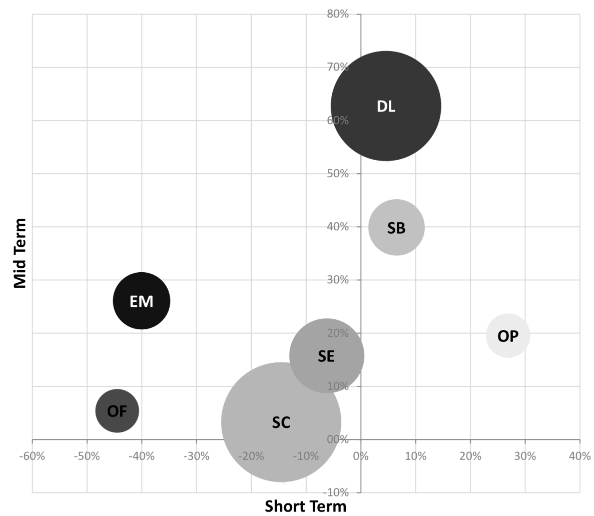

Figure 4: Short- and medium-term contributions of channel contacts in sales conversions

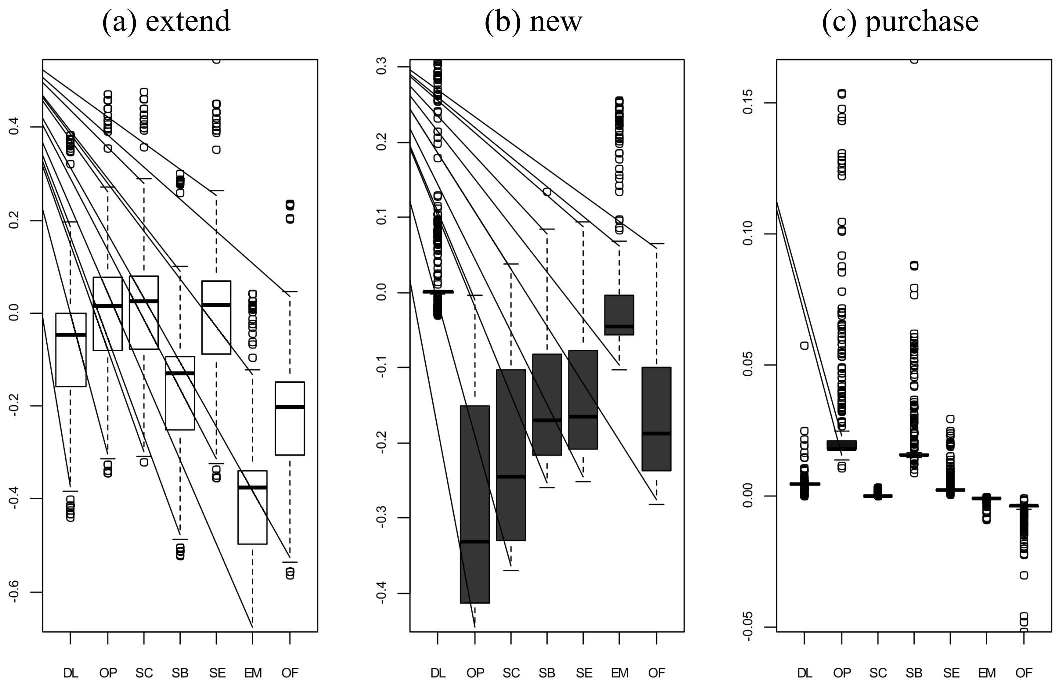

Figure 5: Simulated channel contributions decomposed into extending sessions (panel (a)), returning to initiate a new session (panel (b)) and purchase (panel (c)). (DL is Direct Load, OP is Other Paid, SC is Search Competitive, SB is Search Branded, SE is Shop Engine, EM is Email and OF is Other Free.)