A Dynamic Segmentation Framework: Assessing Omnichannel Behavior of Customers

WORKING PAPER

April 2014

Tanya Mark, Assistant Professor University of Guelph, 50 Stone Road, Guelph, Ontario, Canada, N1G 2W1, 519-824-4120, markt@uoguelph.ca.
Jan Bulla, Associate Professor, Université de Caen, Bd Maréchal Juin, BP 5186
14032 Caen Cedex, France, bulla@math.unicaen.fr.
Rakesh Niraj, Associate Professor, Case Western Reserve University, Cleveland, Ohio, U.S.A., 216-368-0799, rkn10@case.edu.
Ingo Bulla, Institut für Mathematik und Informatik, Universität Greifswald, Walther-Rathenau-Straße 47, 17487 Greifswald, Germany
Theoretical Biology and Biophysics, Group T-6, Los Alamos National Laboratory, Los Alamos, New Mexico, USA
+1 505 665 7670, ingobulla@gmail.com

A Dynamic Segmentation Framework: Assessing Omnichannel Behavior of Customers

Firms are experiencing a dramatic shift in the buying habits of their consumers. In 2010, Statistics Canada reported that 51% of Internet users made purchases online valued at approximately $15.3 billion. MasterCard Canada has also documented this change, noting that while the retail industry had only a modest sales growth of 2.2% during the first six months of 2012,[1] online sales experienced a dramatic growth of 24.9%. Not surprisingly, they emphasize the importance of e-commerce in the retail industry where sales have "increased above the 20% threshold for the ninth consecutive month." Forrester Research reported a similar pattern in the United States. They claim overall retail sales experienced an annual growth rate of three percent between 2006 and 2011while online sales grew by fifteen percent in 2006 and twenty percent in 2011. This growth in online sales represents 46 percent of all US retail sales growth in 2011 [2].

From a firm's perspective, an omnichannel approach to managing customers is important because the firm's priority is to capture as many of a customer's purchase occasions as possible[3]. Allowing customers to choose any channel to both shop and buy from also provides a firm with an opportunity to delight customers (Neslin and Shankar 2009). To successfully deliver an omnichannel approach, firms need to utilize readily available transaction data as well as responsiveness to marketing campaigns. Neslin and Shankar (2009) propose a multichannel customer management decision (MCMD) framework where the first step is to analyze customers with an appropriate segmentation strategy that incorporates channel usage, preference,

---

[1] MasterCard Canada SpendingPulse (TM) Report, August 13th, 2012.
[2] McKinsey&Company, Telecom, Media & High Tech Extranet A symphony of separate instruments: Cross-channel and online sales, Oct. 2012.
[3] The Wall Street Journal, "After Decades of Toil, Web Sales Remain Small for Many Retailers", Shelly Banjo and Paul Ziobro, Aug. 2013.

responsiveness or a combination of these while accounting for changes that may (or may not) occur over time. The segmentation strategy guides the development, design, implementation, and evaluation of the MCMD framework.

The importance of a multichannel segmentation approach has attracted many scholars to this stream of research. The multichannel literature has focused on two main approaches to segmenting customers. First, from the perspective of a single firm, research investigates how to best segment customers based on channel usage. This stream of research defines a multichannel customer as a customer that makes purchases from more than one channel in a given time period and requires longitudinal transaction data (Kumar and Venkatesan 2005; Venkatesan, Kumar, and Ravishanker 2007; Ansari, Mela and Neslin 2008; Valentini, Montaguiti, and Neslin 2011).

The second approach to multichannel segmentation is from the customer's perspective. That is, given a product category, a customer is considered multichannel if a customer purchases products from that category across more than one channel (Kushwaha and Shanker 2013). This perspective derives insight from cross-sectional survey data to understand customer's preferences for each of the channels across product categories (Kushwaha and Shankar 2013) and for both search and purchase phases of the decision process (Verhoef, Neslin, and Vroomen 2007; Konuş, Verhoef and Neslin 2008).

While these scholars have contributed valuable insight into segmenting customers in a multichannel environment, no research has been conducted to understand how a firm can best segment customers beyond characterizations of single versus multichannel behavior (one exception is Valentini, Montaguiti, and Neslin 2011). We address this gap in the multichannel customer management literature by developing a dynamic segmentation framework that is based on a customer's channel usage *patterns* and the evolution of this pattern over time. We argue that customers will increase their usage of the Internet channel over time when they perceive the technology as useful and ease of use is high (Venkatesh et al. 2003). Once the Internet becomes part of the consideration set of channel choice (e.g., web, store, telephone, etc.), customers will vary in their channel usage over time. As a result, different channel usage patterns will emerge because customers use channels for various reasons including inertia, convenience, service, and the influence of marketing activities (Neslin and Shankar 2009). Therefore, we extend prior research by developing a dynamic segmentation framework that accounts for the evolution of channel usage patterns. In our framework, we propose a dynamic model that will estimate a customer's purchasing behaviour conditional on channel choice. Specifically, we extend the hurdle Poisson model by allowing for a multinomial decision in the first part of the model to accommodate the various channels available to customers. In addition, we introduce a hidden Markov model to account for customer heterogeneity and dynamics in responses to a marketing activity (i.e., catalogs) that occur over time. Our approach to capturing patterns in channel usage over time also allows us to address a gap in the multichannel literature raised by Verhoef et al. 2007, "Moreover, if one considers channel choice for purchase, a subsequent interesting point is the quantity purchased in that channel." (page 145).

A second objective of our research is to provide insight into the underlying mechanisms explaining the value of multichannel customers to a firm. Neslin and Shankar (2009) write, "The empirical evidence that the average multichannel customer buys more and is more valuable than the single channel customer is reaching the point of an empirical generalization." (page 72). While there is mounting evidence suggesting multichannel customers are more profitable, little research has been conducted to explain why these customers are more valuable to a firm. In this research, we utilize longitudinal transaction data and a new approach to segmenting customers to assess the underlying mechanisms to explain the monetary value of multichannel customers.

Based on the extant literature, we test for three underlying mechanisms. The first mechanism we investigate is the cost of a channel a customer chooses to buy from. Reflecting the cost variations across channels (e.g., retail store versus Internet), influencing customers to buy from the Internet rather than from a retail store, for example, should improve the overall value of the customer to the company as long as the migration to another channel does not have a negative effect on how much a customer buys. A second mechanism we investigate is whether a retailer can successfully capture more purchase occasions from a customer. With an omnichannel approach to retailing, customers have easier and more convenient access to their favourite retailer. This access could potentially result in customers buying more from the retailer. The third mechanism affecting the profitability of multichannel customers is their price sensitivity. Research suggests that customers that buy online tend to be more deal-prone than customers that buy from traditional channels (Konuş, Verhoef, and Neslin 2008). We formally test for differences in deal-proneness between single and multichannel customers using transactional data.

The remainder of the paper is organized as follows. In the first section, we review the extant literature to understand the limitations of existing segmentation approaches and explicate how the proposed dynamic segmentation framework contributes to this growing body of literature. In the second section, we describe our methodology and the data used to empirically validate our dynamic segmentation framework. The third section presents our results and we conclude with a discussion of the contributions of our research, limitations, and directions for future research.

Literature Review

Dynamics in Channel Choice over Time

Early research in the multichannel customer management literature investigated the impact of marketing activities on a customer's migration from traditional channels to the Internet (Ansari, Mela, and Neslin 2008), and also focused on understanding the characteristics and perceptions of customers buying from single channels versus multiple channels (Konuş, Verhoef and Neslin 2008). Researchers have developed sophisticated models to predict cross-channel elasticities (Avery, Steenburgh, Deighton, and Caravella 2012) and the allocation of resources across marketing activities and channels (Wiesel, Pauwels, and Arts 2011). The central notion underlying research on multichannel customer management is how to improve the customer experience across channels over the duration of the customer-firm relationship (Neslin et al. 2006). By exceeding customers' expectations at each touch point, firms will be able to strengthen relationships with their customers and ultimately increase the overall value of their customer base.

Valentini et al. (2011) develop a dynamic model to explain channel preference as a function of state dependence and marketing activities. They assume only two stages exist in the channel choice process, namely trial and post-trial stages and estimate two separate logit models, one for each stage, to predict channel choice during that stage. Although they find support for the evolution of channel choice over time, they argue future research is needed to investigate whether additional stages exist in the customer channel choice process. To the best of our knowledge, no study has explicitly accounted for the evolution, or dynamics, in channel usage over time while accounting for the different types of relationships between customers and firms.

We fill this research gap identified by them. See Table 1 for a summary of the key findings in the multichannel literature.

INSERT TABLE 1 ABOUT HERE

While we will focus on managing customer's relationship beyond the first purchase across multiple *purchase* channels, more recently, managing the communication process across multiple *communication* channels through various stages in the buying process leading up to a purchase has also been examined (Konuş, Verhoef, and Neslin 2008; Valentini, Montaguti, and Neslin 2011; Abhishek, Fader, and Hosanagar 2012). Konuş et al. model the total utility of a shopping experience by considering both product information search, and the product purchase processes. They find multiple segments that differ due to consumer preference for using multiple channels versus a single channel for both search and purchase stages in the buying process. A limitation of their study is the use of cross-sectional data to uncover stages in a dynamic process.

Abhishek, Fader, and Hosanagar (2012) develop an attribution model to investigate the effectiveness of an ad campaign at migrating customers along the conversion funnel leading to a purchase. Using a hidden Markov model in an automotive context, they consider the entire path a consumer takes from a dormant state to a state of consideration rather than only considering the last impression before the consideration state. As a result, the model assigns less credit to display impressions on generic websites and more credit to display impressions when there is congruency between the impression and a specific website. This substantive finding would not be possible without a dynamic model that captures the evolution of the decision process.

As such, we consider a similar approach to investigate how purchasing behavior evolves over time in multichannel context while accounting for channel usage and marketing responsiveness. Our approach uses longitudinal transaction data that spans several channels including retail store, Internet, catalog, and call center. This unique dataset enables us to estimate latent states based on purchasing behavior whereas they use an advertising agency's data on web browsing activity during an ad campaign. Furthermore, we capture dynamics of customer-firm relationships across multiple channels and therefore provide a richer conceptualization of the evolving relationship between a firm and customers. This conceptualization is consistent with the customer-firm relationship literature (Netzer et al., 2008; Montoya et al. 2010; Ascarza and Hardie 2012; Mark et al. 2013) that identifies multiple states and also underscores the need to account for relationship dynamics when assessing purchasing behavior.

Impact of Direct Mailings on Purchasing Behavior for an Omnichannel Retailer

As consumers develop new habits with various channels available for them to both shop and buy from, understanding how responses to marketing activities vary by channel is managerially relevant and also important to academics conducting research on multichannel behaviour. Neslin and Shankar (2009) argue, "Segments formed on the basis of channel are probably differentially responsive in that customers who use different channels most likely have different needs (e.g., for convenience versus service) and therefore users of Channel A will respond differently to a given marketing action than users of Channel B" (page 71). Although this area of research is new in the field of marketing (Dholakia et al. 2009), there are seminal studies that have contributed toward our understanding of individual differences in consumer responses to marketing activities across channels.

Ansari, Mela and Neslin (2008) find heterogeneity in the effects of catalogs on channel migration. Specifically, most customers were more likely to buy from the catalog channel (i.e., call centre) after receiving a catalog. However, for some customers catalogs had a significant negative effect on the catalog channel suggesting catalogs are also an effective marketing communication to influence customers to migrate to the Internet channel. In their seminal article, they also found congruency between the channel and marketing activity (i.e., catalog versus email promotion) increases the likelihood of a purchase occasion in that particular channel.

This stream of research has also investigated cross-channel effects of *communication* activities on *purchasing* behavior. Naik and Peters (2009) develop an integrated marketing communications model to investigate the interaction of offline and online media. Their hierarchical model identifies and quantifies synergies from both types of advertising, though not for all media. Similarly, Dinner, Van Heerde, and Neslin (2011) estimate *purchases* in both online and offline channels as a function of online and traditional *communication* activities. Their findings suggest that traditional advertising is driving transactions that might have occurred online to the stores, cannibalizing online customer counts.

More recently, Neslin et al. (2013) develop an optimization model to maximize the profitability of customers using a customer's purchase recency and responses to both offline and online communications. They define 216,000 states based on predefined recency states for each month and several lagged marketing states. They find optimal levels of direct mail are higher than those of email as direct mail has stronger carryover effects while email has a higher saturation effect. Building upon their contributions in incorporating recency of purchases, we propose a more parsimonious model, which we describe fully later in the methodology section, and which simplifies the number of states in a company's customer base tracked over time.

Mechanisms Explaining the Value of Multichannel Customers

The multichannel literature offers mounting evidence in favour of encouraging consumers to buy from more than one channel (e.g., Internet and retail store) since it is linked to an increase in the number of purchases (Ansari et al. 2008) and more profitable customers (Venkatesan, Kumar, and Ravishanker 2007). However, Ansari and his co-authors warn marketers: "the notion that migration is unqualifiedly positive because it lowers costs and increases demand should be tempered by the admonition that it can be negatively associated with long-term purchase patterns." (page 71). They posit that lower switching costs and lack of interaction with other humans reduces a customer's loyalty. Indeed, Konuş et al.'s (2008) findings further support this argument that customers with multichannel preferences tend to be less loyal to any particular brand and are more price conscious than customers with single channel preferences. Campbell and Frei (2010) find in a banking context that desirable outcomes like satisfaction, retention and market share increases while profitability was lower for customers using multiple channels, primarily for cost reasons. These contradictory findings on the value of the Internet as an additional channel suggest the Internet has an important role in contributing to a multichannel customer management strategy but more research is needed to disentangle the long-term effects of marketing activities and channel choice across channels on buying behavior.

From a customer's perspective, Kushwaha and Shankar's (2013) examine the moderating role of product characteristics on the relationship between channel preference and purchase expenditure. They define multichannel customers as customers who make purchases within one broad category from more than one channel. For example, a consumer might purchase a pair of shoes from a retail store and buys an additional pair of shoes over the Web at the next purchase occasion. This channel usage pattern contrasts with a single channel customer who consistently purchases shoes offline from various retail stores. According to Kushwaha and Shankar, multichannel customers spend more within the category, but not necessarily at a given retailer, than single-channel customers regardless of the product categories.

Given the mixed findings in the literature, we further investigate the relationship between multichannel behavior and profitability by assessing the underlying mechanisms explaining the value of multichannel behavior. In this research, we take the perspective of a single firm, albeit with multiple channels. We consider three mechanisms that have been put forward by prior literature to explain why multichannel behavior may be more valuable to a retailer. First, we explicitly account for the cost of serving customers by channel. Each channel varies in its costs, either directly or indirectly, which therefore affects the overall profitability of serving a customer across different channels. As such, the first mechanism explaining the value of multichannel customers is the differential cost required to serve customers based on their channel usage.

Second, we argue that a retailer with multiple channels can better serve its customers than a single channel retailer because of a better potential fit between a customer's need for convenience, accessibility, and other features to what the multiple channels can offer. Because of these factors and findings in the channel migration literature (Ansari, Mela, and Neslin 2008), a retailer with multiple channels will gain a larger portion of a customer's total purchase occasions. Individuals who buy using multiple channels offered by a retailer are likely to be buying across multiple types of needs and situations. For example, they may be going to a store to buy for themselves, but may use the Internet channel for gift giving. Thus, we posit that a second mechanism to explain the value of multichannel customers is the growth in the total number of purchases a customer makes with a particular retailer.

Third, literature suggests that customers who purchase from multiple-channels also tend to be deal-prone (Konuş et al. 2008; Neslin and Grantham 2013). Price conscious customers are less valuable to a retailer than non-price conscious customers. Therefore, our third mechanism is whether or not a customer exhibits price conscious tendencies when making purchases from a retailer. In this research, we investigate these three mechanisms empirically.

Data

We collected individual-level transaction data from a major North American retailer that sells apparel and household goods across various channels, including retail stores, telephone, mail, and the Internet[4]. Customers for our data were selected as a random sample from all the customers of the retailer who made a minimum of three purchases over the duration of the nine-year observation period. We selected a sample of 1,000 customers, and observed them over a nine-year period, beginning in January 2001. We divided the observations into thirty-six quarters. This provides for a long series for the dynamics while also keeping the data manageable. Since the items customers purchase are not characterized by very high frequency (i.e., median inter-purchase time is 2 months), aggregating the data to quarterly intervals also reduces zero-inflation in the dataset. The firm also makes marketing decisions and tracks performance in quarterly cycles, so it is a natural time-period from a managerial perspective.

[4] For confidentiality reasons, the retailer has requested to remain anonymous. The dollar values are multiplied by a factor so as to disguise actual values. In this paper, the mail channel is negligible and hence it is combined in the telephone channel henceforth.

Aggregating the data to quarterly is also consistent with other research modeling the evolution of channel choice over time in a retail context (Valentini, Montaguti, and Neslin 2011).

Each customer in this cohort made her first purchase from the retailer during the first year of the observation period. It may be noted that the nine-year period beginning in 2001 is characterized by increasing adoption of Internet as a buying channel for customers in North America, as also evident in our sample. Our sample is predominantly female (81% female, 19% male) and 28% of them are not married. Conditioned on making a purchase, they have a mean number of orders of 1.85 with an average purchase value of $173.99 per quarter. A total of 29.2% of customer/quarters see no purchases, indicating zero-inflation. More precisely, our dependent variable exhibits substantial overdispersion: We test for equidispersion and for the appropriateness of the simple Poisson model (i.e., H_0: Y ~ Poisson (λ) against H_1: var(Y) > λ) using the T statistics of Böhning (1994) and refinend by Baksh, Böhning, and Lerdsuwansri (2011). Both tests reject the hypothesis of equidispersion in favor of overdisperison due to zero-inflation (p-value = 0).

Out of the remaining 70.09% active customer/quarters most are characterized by purchases from a single channel: most from the telephone channel (43.79 %), followed by the Internet (10.28 %) and the retail store (3.74 %). We also captured multichannel behavior that we define as more than one purchase by a customer across two or more channels within a quarter. We find 12.28% of customer/quarters can be characterised as multichannel. Figure 1 is a graphical representation of the time trends of purchasing activity and channel usage over the 9-year period. The time trends clearly depict seasonality around the holiday season. We account for this seasonality in our HMM model by incorporating a dummy variable for the last quarter of each year in our data. The channel usage trend demonstrates a significant decline in the telephone channel, while the other channels increase slightly by the end of the observation period. In addition, we observe the across channel usage is cyclical with peaks during the holiday season.

Our dataset also includes information on direct mailings sent to each customer. The retailer did not use any other marketing activity during the observation period and had consistent pricing across all channels. Over the entire period and across all customers in our sample, a customer receives a mean of 13.12 catalogs per quarter. In Figure 2, the time trend for the number of catalogs mailed to customers suggests the retailer sends a larger number of catalogs to customers earlier in the customer-firm relationship compared to later in the relationship. In addition to channel choice and a marketing instrument, we included demographics (i.e., gender and whether or not an individual was married) and several dummy variables to represent the last quarter of each year to estimate our HMM. These covariates were included to account for additional heterogeneity present in our sample as well as for the seasonality around the holiday season. Drawing from research on dynamic catalog mailings (Simester, Sun, and Tsitsiklis 2006), we also included a lagged catalog variable in our model to account for the effect of a catalog beyond a customer's current purchase occasion.

INSERT FIGURES 1 & 2 ABOUT HERE

Methodology

The Hurdle-Poisson Hidden Markov Model

We begin this section by describing our dynamic segmentation framework using a hurdle-Poisson hidden Markov model. This section is followed by a description of our post hoc analysis. As described in the previous section, our data are subject to zero-inflation and overdispersion. Additionally, effects such as time dynamics and effects of covariates on channel choice and purchase frequency may be present and vary among individuals. Therefore, we propose a hidden Markov model that will estimate a customer's purchasing behavior conditional on channel choice. Hidden Markov models have been applied to various contexts in the customer management literature (Netzer et al. 2008; Montoya et al. 2010; Abhishek, Fader, and Hosanagar; Mark et al. 2013). This stream of research suggests purchasing models are richer and provide more accurate parameter estimates when they account for the dynamics inherent in the decision process and customer-firm relationship.

Our model extends the Poisson hurdle model by 1) allowing for a multinomial decision in the first part of the model, and 2) introducing random coefficients driven by a latent Markov chain. It may be noted that the model can be extended to incorporate a large class of parametric distributions other than the Poisson. However, in the following we restrict our explanations to the Poisson hurdle model to avoid unnecessary notational confusion.

In a basic HMM for longitudinal data, the existence of two processes is assumed: an unobservable finite-state first-order Markov chain, S_{it}, i= 1, …, n, t= 1, …, T, with state space S= {1, …, m} and an observed process, $Y_{it,}$ where Y_{it} denotes the response variable for individual *i* at time *t*. The distribution of Y_{it} depends only on S_{it}, specifically the Y_{it} are conditionally independent given the S_{it}. However, without this conditioning the Y_{it} are not independent in time. Thus, the unknown parameters in a HMM involve both the parameters of the Markov chain and the state-dependent distributions of the random variables Y_{it}. In particular the parameters of the Markov chain are the transition probabilities $Q = \{q_{itjk}\}$, where $q_{itjk} = Pr(S_{it} = k|S_{it-1} = j)$,j,k

$\in$S is the probability that individual *i* visit state *k* at time *t* given that at time *t-1* she was in state *j*, and the initial probabilities $\delta_{ij} = \Pr(S_{i1} = j)$, i.e. the probability of being in state *j* at time 1.

In the statistics literature, the hurdle model has been proposed as an appropriate approach for the analysis of longitudinal counts with inflated or deflated zero counts (Todem et al. 2010). The hurdle model is a two-part conditional model where the first part of the model consists of a point mass at zero, referred to as the hurdle, which we extend to a 5-dimensional vector. This vector consists of four components that represent a purchase occasion through one of the four channels, including multiple channels in any given period, and the fifth is the no purchase event. The second part of the model is a truncated Poisson (or over-dispersed truncated Poisson) distribution to model the frequency of purchases by channel.

To account for a time varying association structure, we relax the independence assumption between the processes made by the basic hurdle model and introduce a common latent structure, assumed to follow a Markov chain. The Markov chain represents the unobservable heterogeneity in purchase incidence and frequency by channel. Formally, let y_{it} be the observed number of orders, k_{it} the channel variable taking values k from 1 to K and $d_{it} = (d_{it}^1, \ldots, d_{it}^K) = (I(k_{it} = 1), I(k_{it} = 2), \ldots, I(k_{it} = K))$ a K-dimensional dummy vector indicating the channel selected. Here, the channel K represents no purchase, that is $I(k_{it} = K)$ is equivalent to $y_{it} = 0$. Conditional on the hidden state, $S_{it} = j$, the vector d_{it} follows a multinomial distribution with canonical parameters, $[\pi_{itj}^1, \ldots, \pi_{itj}^K]$ modeled through the multinomial logit link, i.e.,

$$\pi_{itj}^k = \frac{e^{v_{it}'\phi_j^k}}{1+\sum_{l=1}^{K-1} e^{v_{it}'\phi_j^l}} \quad \text{for} \quad k \in \{1, \ldots, K-1\} \text{ and, naturally,}$$

$$\pi_{itj}^{K} = \frac{1}{1 + \sum_{l=1}^{K-1} e^{v_{it}'\phi_j^l}}$$

where $v_{it} = \left[(v_{it})_1, (v_{it})_2 \dots, (v_{it})_q\right]$ represents a q-dimensional covariates vector with intercept $(v_{it})_1 = 1$ and $\phi_j^k = \left[\left(\phi_j^k\right)_1, \left(\phi_j^k\right)_2 \dots, \left(\phi_j^k\right)_q\right]$ are the corresponding state-specific parameter vectors for the channels 1 to K-1. Similarly the positive count process is assumed to follow a truncated Poisson distribution, conditionally on $S_{it} = j$

$$\frac{f(y_{it}; \lambda_{itj}^k | x_{it}, S_{it} = j)}{1 - f(0; \lambda_{itj}^k | x_{it}, S_{it} = j)} = \frac{\frac{\left(\lambda_{itj}^k\right)^{y_{it}}}{y_{it}!} e^{-\lambda_{itj}^k}}{1 - \frac{e^{-\lambda_{itj}^k}}{y_{it}!}}$$

where the canonical parameter λ_{itj}^k is modeled in a generalized linear model framework as

$$\log\left(\lambda_{itj}^k\right) = x_{it}'\beta_j^k$$

with $\beta_j^k = \left[\left(\beta_j^k\right)_1, \left(\beta_j^k\right)_2 \dots, \left(\beta_j^k\right)_p\right]$ the vector of state-specific fixed parameters associated with the covariate vector $x_{it} = \left[(x_{it})_1, (x_{it})_2 \dots, (x_{it})_p\right]$ with intercept $(x_{it})_1 = 1$ for channel k, where $k \in \{1, \dots, K-1\}$. Note that the choice of f(·) may also include, among others, log-normal, negative binomial as well as other parametric distributions. Evidently, the link-function then has to be adapted to the new situation.

Under the model assumptions described above, the likelihood function may be represented in a convenient form (Zucchini and MacDonald, 2009):

$$L(\boldsymbol{\theta}) = \prod_{\boldsymbol{i=1}}^{\boldsymbol{n}} (\boldsymbol{\delta}_i \boldsymbol{P}(y_{i1}) \boldsymbol{Q} \boldsymbol{P}(y_{i2}) \boldsymbol{Q} \dots \boldsymbol{P}(y_{iT-1}) \boldsymbol{Q} \boldsymbol{P}(y_{iT}) \mathbf{1}'),$$

where $\boldsymbol{P}(y_{it})$ represents a diagonal matrix with the state-dependent conditional distributions as entries, i.e.,

$$\boldsymbol{P}(y_{it}) = \begin{pmatrix} \left(\pi_{it1}^{K}\right)^{d_{it}^{K}} \prod_{k=1}^{K-1} \left[\pi_{it1}^{k} \frac{f\left(y_{it}; \lambda_{its_{it}}^{(k)} \middle| x_{it}, S_{it} = 1\right)}{1 - f\left(0; \lambda_{its_{it}}^{(k)} \middle| x_{it}, S_{it} = 1\right)} \right]^{d_{it}^{k}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \left(\pi_{itm}^{K}\right)^{d_{it}^{K}} \prod_{k=1}^{K-1} \left[\pi_{itm}^{k} \frac{f\left(y_{it}; \lambda_{its_{it}}^{(k)} \middle| x_{it}, S_{it} = m\right)}{1 - f\left(0; \lambda_{its_{it}}^{(k)} \middle| x_{it}, S_{it} = m\right)} \right]^{d_{i1}^{k}} \end{pmatrix}$$

Moreover, $\boldsymbol{Q}$ represents the transition probability matrix and $\boldsymbol{\delta}_i$ a vector with the initial probabilities as entries. The two most popular approaches for maximizing the likelihood function are, on the one hand, the so-called expectation–maximization (EM) algorithm presented to a larger public by Dempster et al. (1970). On the other hand, direct optimization of the likelihood, e.g. by quasi-Newton algorithms, may be utilized and constitutes the second technique preferred by a larger community. Both methods have strengths and weaknesses, which is why we utilize a hybrid algorithm as presented by Bulla & Berzel (2008) where the algorithm starts with an optimization by means of the Nelder-Mead algorithm. After a certain number of iterations, it switches to quasi-Newton optimization until full convergence is achieved. In order to ensure that the final solution is a global and not local maximum, various sets of random initial values as well as an initialization by a non-informative prior are usually carried out. Unfortunately, the Hessian matrix, which is obtained as by-product from the quasi-Newton algorithm, does not provide numerically stable results for long time series. Therefore, standard errors of the parameter estimated have to be computed by a parametric bootstrap approach (Bulla & Berzel 2008).

Post Hoc Analysis

We perform post hoc analysis to investigate dynamics in multichannel behavior over time. We use the HMM a posteriori classification probabilities to determine state membership for each customer at each time period. We compute the posterior probabilities of a customer being classified to a state for each time period. Using a maximum a posteriori analysis of the posterior probabilities, we then determine state membership for each customer for each quarter. These state classification probabilities can inform us on dynamics in purchasing behavior across a retailer's channels over time. In this paper, we argue that the states represent different types of relationships between customers and firms. This assumption is consistent with other dynamic models in the customer relationship management literature (Netzer et al., 2008; Montoya et al. 2010; Mark et al. 2013). Furthermore, our segmentation approach enables a deeper insight into how channel preference changes over time as a function of the different types of customer-firm relationships.

Results

To investigate channel usage patterns over time, we compare our proposed dynamic model of channel choice to traditional models in the marketing literature. The first model we compare our model to is a hurdle model with no dynamics. The second comparison model is a latent class model with varying number of latent classes, i.e., customers do not transition among the classes. We refer to the Bayesian Information Criteria (BIC) and our preference for a parsimonious model, and find a four state hidden Markov model (HMM) is superior to traditional models found in the marketing literature. Refer to Table 2 for a summary of the BIC for each model.

The estimates of the HMM parameters provide insight into the initial probabilities of a customer belonging to a state as well as the likelihood of transitioning from one state to another over time (see Table 3 for HMM parameters). Thus a majority of the customers in the sample begin in state 1 (84%). Almost fourteen percent begin in state 2 and the remaining two percent begin in states 3 and 4 with a larger portion starting in state 4 (1.8%). The transition probability matrix suggests that state membership is highly persistent for each of the four states suggesting that if a customer transitions to another state, she is likely to remain in that state for the duration of the observation period.

INSERT TABLES 2 & 3 ABOUT HERE

Second, we compute the posterior probabilities of a customer being classified to a state for each time period. We find the largest number of transitions occurs for customers who begin in state 1. These customers are more likely to transition to state 4, followed by states 2, and 3 respectively. Customers in state 2 are most likely to transition to state 3. State 3 customers rarely transition to any other state so this seems to be the most absorbent state. Finally, state 4 customers are likely to transition to state 2 followed by state 1. Few of these customers transition to state 3. We also examined the number of transitions based on the posterior probabilities. We find 22.30% of the customers do not transition to another state. Almost half of the customers make one transition (48.90%) and 21.10% make two transitions. Less than 8% of the customers make more than three transitions over time. See Table 4 for a summary of the inferred transitions and Table 5 for a summary of observed customer dynamics.

INSERT TABLES 4 & 5 ABOUT HERE

We begin by examining our HMM model intercepts. The first step of the model, the multinomial choice, estimates the probability of choosing a channel, conditional on a purchase incidence in the quarter (see Table 6 for the multinomial choice parameter estimates). We find the first state can be characterized as the traditional catalog customers because they have a higher propensity to use the telephone to place an order [$(\varphi_j^k)_1$= (φ_1^1=-5.816, φ_1^2=-1.825, φ_1^3=-6.275, φ_1^4=-5.682)]. State 2 consists of customers that have equal affinity for physical store purchases and across channels [$(\varphi_j^k)_1$= (φ_2^1=.622, φ_2^2=-1.293, φ_2^3=-3.092, φ_2^4=-1.397)]. We empirically investigate these results and find state 2 customers have a much higher proportion of multichannel purchases compared to customers belonging to the other three states. Specifically, customers in state 2 make, on average, 62% of their purchases from multiple channels compared to 13.12%, 14.37% and 9.61% for states 1, 3, and 4 respectively (pairwise t-test is significant at p < .05). Although each state has customers that buy from more than one channel over the 36 quarters, we argue that state 2 customers represent the multichannel customers because this buying behavior is consistent over time. Compared to others, state 3 customers are mostly inactive customers; however, when they make a purchase, they are most likely to purchase from the retail store [$(\varphi_j^k)_1$= (φ_3^1=-3.252, φ_3^2=-5.625, φ_3^3=-6.302, φ_3^4=-10.033)]. Finally, state 4 customers are primarily the web-purchasing customers [$(\varphi_j^k)_1$= (φ_4^1=-2.347, φ_4^2=-1.926, φ_4^3=-.543, φ_4^4=-5.137)].

INSERT TABLE 6 ABOUT HERE

Now we examine the evolution of state membership. Using the classification probabilities, we can determine the evolution of the latent states over time. We find all of the states have an increase in the number of customers classified into the respective state, except for state 1 which clearly has a decreasing trend (See Figure 3). We confirm these trends using a simple linear regression on the proportion of customers in each state on time and find all four states have a highly significant slope ($p < 0.01$).

INSERT FIGURE 3 ABOUT HERE

Differential Impact of Direct Mailings on Purchasing Behaviors

We now investigate how a marketing activity affects purchasing behaviors over time. Recall that our proposed HMM model is a two-part model: first, we estimate the probability of choosing a particular channel (multinomial choice) and second, the frequency of purchases by channel ($y > 0$, i.e., truncated Poisson) conditional on a customer making at least one purchase in the quarter. Refer to Table 6 for the multinomial choice model parameter estimates and Table 7 for the truncated Poisson model parameter estimates. Subsequently, we describe the multinomial choice parameter vectors followed by the truncated Poisson parameter vector for each of the four states.

Multinomial Choice Parameter Estimates

For state 1, we find the number of catalogs received by a customer has a significant effect on purchase incidence across all four channels after controlling for demographics and seasonality [$(\varphi_j^k)_4$= (φ_1^1=.076, φ_1^2=.125, φ_1^3=.046, φ_1^4=.167)]. This marketing instrument has the strongest impact on multichannel behavior for customers belonging to state 1 implying these customers are more likely to make an additional purchase from several channels if they receive catalog. The lag of catalogs did not have a significant effect on purchase incidence for any of the channels, which suggests the effect of catalogs may not extend beyond a three month period for these customers. For customers classified in state 2, catalogs have a positive significant effect across each of the channels except for the retail store [$(\varphi_j^k)_4$= (φ_2^1=-.089, φ_2^2=.063, φ_2^3=.108, φ_2^4=.100)]. The variable representing the lagged effect of catalogs has a negative effect on retail store purchases but is not significant for any other channel [$(\varphi_j^k)_5$= (φ_2^1=-.037)]. For state 2 customers, catalogs are more likely to increase the likelihood of a customer buying from either the web or across channels, suggesting they are more likely to make more than one purchase across channels after receiving a catalog in one quarter. Similar results are found for state 3 customers. Catalogs have a significant effect on purchase incidence for all of the channels except for the retail store [$(\varphi_j^k)_4$= (φ_3^2=.243, φ_3^3=.400, φ_3^4=.528)]. The lag of catalog promotions is more likely to influence purchase incidence for both the web and multichannel behavior but is not significant for the other channels [$(\varphi_j^k)_5$= (φ_3^3=.185, φ_3^4=.219)]. These findings suggest customers belonging to state 2 and 3 are more likely to buy from channels other than the retail store as a result of receiving an additional catalog.

In state 4, we find a different pattern. For this state, catalogs have negative impact on purchase incidence for both retail store and Web purchases, but they increase the likelihood of purchase over the telephone and across channels [$(\varphi_j^k)_4$= (φ_4^1=-.072, φ_4^2=.256, φ_4^3=-1.085, φ_4^4=.275)]. The lag of catalogs has similar impact on purchase incidence but is not significant for phone purchases [$(\varphi_j^k)_5$= (φ_4^1=-.117, φ_4^3=-.886, φ_4^4=.058)]. Overall, our findings suggest catalogs are more likely to increase the likelihood of purchase incidence over the telephone.

Truncated Poisson Parameter Estimates

For the second part of our model, we find important differences between purchase incidence and frequency for each state. Notably, catalogs have a significant negative impact on frequency of purchases over the Web for customers classified in state 1 but continue to have a positive impact on store purchases but is not significant for phone and across channels [$\left(\beta_j^k\right)_4$= (β_1^1=.120, β_1^3=-.232)]. State 2 customers, those that are likely to be multichannel, are more likely to increase the number of purchases over the Web after receiving a catalog but the effect is not significant for any other channel [$\left(\beta_j^k\right)_4$= (β_2^3=.028)]. State 3 customers are the only customers that catalogs have similar effects across channels for both purchase incidence and frequency of purchases. For state 4 customers, catalogs have a positive impact on the frequency of purchases at the store and over the phone but a negative impact on Web purchases [$\left(\beta_j^k\right)_4$= (β_4^1=.186, β_4^2=.060, β_4^3=-.215)]. The lagged effect of catalogs was not significant for states 1 and 2, but was statistically significant for states 3 and 4. The effect is negative for the phone channel but positive for the Web for state 3 customers [$\left(\beta_j^k\right)_5$= (β_3^2=-.293, β_3^3=.050)]. For state 4, the lagged variable had a significant negative effect on both the frequency of purchases in the store and web channels [$\left(\beta_j^k\right)_5$= (β_4^1=-.210, β_4^3=-.115)]. Thus, for customers belonging to states 3 and 4, the effect of catalogs is significant beyond the immediate purchase occasion. Together, these findings suggest that catalogs will differentially impact channel purchase incidence and frequency based on the type of customer-firm relationship that exists for each of the four states.

INSERT TABLE 7 ABOUT HERE

Mechanisms Explaining the Value of Multichannel Customers

To investigate the underlying mechanisms explaining the value of multichannel customers, we conducted several statistical tests to determine if there is a significant difference in value between multichannel and single channel customers. Recall from our discussion in the previous section that state 2 is the state characterized by multichannel behavior. However, also note that over the 9-year period, customers transition from state to state. Thus, in these comparisons, we take all the customer/quarters where a customer is inferred to be in state 2 based on our post hoc classification and compare that with all the customer/quarters where a customer is in any other state. We began this analysis by first consulting with the management team of the company to determine the relative contribution margins for each of the channels. We were informed the retail channel has the lowest contribution with an average contribution margin of around 30%, compared to 40% for telephone, and 50% for the Web. Using these margins, we tested for significant differences in profitability between multichannel and single channel customers and find that multichannel customers have higher margins than others ($t = -13.84$, $p < .05$). We also observed, in the quarters with multichannel purchases, Internet channel was one of the channels 60.7% of the times, while in single-purchase quarters, Internet accounted for only 17.6% of all purchases. These proportions are significantly different at $p < .05$.

We now examine the second mechanism as a possible explanation to why multichannel customers are more valuable to a retailer. Based on the extant literature, we posit that multichannel customers are more valuable to the firm because they tend to buy more than single channel customers. We formally test for differences in revenue between state 2 customers and customers belonging to any of the remaining states and find that multichannel customers spend more with the retailer relative to single channel customers ($t = -16.89$, $p < 0.0001$).

For the third mechanism, we assess price sensitivity of multichannel customers compared to customers that are predominantly single channel customers. In our dataset, the retailer also included a variable that recorded whether or not a purchase was made while on promotion. We use this variable to provide insight into the price sensitivity of multichannel customers. We find on average only 12.78% of customers made a purchase while a product was on sale. We compare the proportion of purchases made by customers in State 2 when a product they bought was reported on-sale for the quarter, to the proportion of purchases made by customers in all other states. We find multichannel customers are not different from other customers ($t=1.09$, p-value $> .10$) and therefore do not buy disproportionately more on promotion. Interestingly, we find customers classified to state 3 have a statistically significantly higher propensity to buy on promotion and they possibly represent a deal-prone segment.

Together, our findings are consistent with the multichannel literature in that multichannel customers are more valuable to a firm because they tend to buy more from a retailer and these customers use the Internet channel which has a higher contribution margin. However, we find multichannel customers are no different from other customers in terms of their price sensitivity.

Discussion

We begin this section with our contributions to research followed by managerial implications. We conclude with suggestions for future research and the limitations of our research.

Our research has several contributions. First, we contribute insight toward understanding the multiple effects of direct mail on purchasing behavior across channels over time. Research suggests congruency between a marketing communication and purchasing channel will increase the effectiveness of the campaign in that particular channel (Ansari et al. 2008) as well as across channel (Naik and Peters 2009; Dinner et al. 2011). Our findings are consistent with this stream of research in that across all four states catalogs increase the likelihood of a purchase in the telephone channel but only when we examine purchase incidence. In our research, we extend the hurdle model to allow for a multinomial choice in the first part of the model thereby enabling insights into not only purchase incidence across channels, but also into the frequency of purchases a customer makes in any given period over an extended period of time. When we assess the impact of catalogs on the frequency of purchases across channels, we find differential effects of catalogs on purchasing behavior for customers belonging to each of the four states. Specifically,

- For state 1, catalogs are more likely to increase the frequency of purchases in the store.
- For state 2, catalogs have a stronger effect on web purchases and are not significant for phone purchases.
- For state 3, these customers make infrequent purchases, however; when they do make a purchase they are more likely to buy from the phone, followed by web and across channels.
- For state 4, customers that predominantly use the Internet to make a purchase, catalogs have the strongest impact on increasing the frequency of purchases in the store.

Together these findings suggest catalogs will increase the number of purchases in the telephone channel, but only for the first purchase incidence and for customers with infrequent buying patterns. For active customers, the catalog is an effective tool for increasing the likelihood of additional purchases across channels. These substantive findings would not be possible without our novel model that simultaneously accounts for customer heterogeneity, channel choice, and responsiveness to a marketing activity. Therefore, our research contributes to our understanding of the multiple effects of catalogs on purchasing behavior beyond the first purchase.

Our second contribution to research is insight into a customer's journey to purchase and beyond. We address one of Marketing Science Institute's top research priorities for 2012 - 2014: "Rethinking the journey to purchase and beyond". Specifically, our research addresses the need for "research that maps how the actions of firms influence buyer states" [5] by developing a dynamic segmentation framework. Our framework offers a richer conceptualization of a customer's buying pattern over time by simultaneously assessing how a marketing activity affects (1) channel choice, (2) frequency of purchases, and (3) the evolution of buyer states over time beyond the first purchase.

At the aggregate level, we observe a decrease over time in usage of the telephone channel and a corresponding increase in usage of the other channels offered by the retailer (see Figure 1). The decreasing trend in the telephone channel usage is not surprising: a new technology (i.e. Internet) has disrupted habits and preferences for existing purchasing channels. This behavioral change is to be expected with disruptive technologies but over time we should observe a new stable and persistent channel usage pattern emerge (Lyytinen and Rose 2003). Indeed, over the

[5] Marketing Science Institute 2012-2014 Research Priorities. (Nov. 2013). Retrieved from http://www.msi.org/research/msi-research-priorities/priority-2-rethinking-the-journey-to-purchase-and-beyond/.

nine-year period, we do find customers adopt the new channel. For some customers it becomes a new preferred channel, for others it augments existing preferences, and yet others remain true to their original channel. Our dynamic model is flexible in that it accounts for each unique pattern a customer chooses throughout the customer-firm relationship. Furthermore, our model segments customers accounting for evolution of channel choice over time rather than segmenting customers based on static channel choices (i.e., single versus multichannel customer). Thus, we accommodate the dynamics in channel choice for each customer recognizing that it is inevitable that customers are going to try a multitude of channels and over time they will establish new habits to achieve their shopping and buying goals. Indeed, in our post-hoc analysis, we find that only a small minority (22.30%) shows no evolution over the four states during their relationship with the firm, and most have one (48.90%), two (21.10%) or even more number of transitions (7.7%). Retailers that offer an omnichannel approach and account for these dynamics in their response models will make better decisions and ultimately will have more effective marketing activities that will increase the overall value of their customer base.

Our third contribution to the literature is insight into the profitability of different types of customer-firm relationships. The customer-firm relationship literature suggests loyal customers are more valuable to a firm because they are more likely to buy more, recommend the firm and its products to others, and are willing to pay a premium for a brand (Rust, Lemon, and Zeithaml 2004). From a channel choice perspective, the literature suggests that multichannel customers tend to be less loyal over time, more price-sensitive, and as a result less profitable to a single retailer (Ansari et al. 2008). To date, no research has been conducted to assess the value of different types of customer-firm relationships and their channel choice. In this research, we address this gap by developing a channel choice model that segments customers based on latent relationships between customers and firms and their responses to a marketing instrument. By doing so, we are able to provide insight into the value of a customer over time based on the strength of the customer-firm relationship as well as her channel choice. Furthermore, we test for three underlying mechanisms to explain the relative value of customers. We find multichannel customers are more valuable to the firm because they buy from a variety of channels and often more from newer, lower cost, Internet channel and these customers tend to buy more than other customers. We did not find support for multichannel customers having a higher propensity to be price sensitive. Our findings are more consistent with the promotion response literature that suggests there is a deal-prone segment (Narasimhan, Neslin, and Sen 1996), regardless of channel choice.

Implications for Retailers

We are in an era where retailers are inundated with data (i.e., Big Data) and marketers are astutely aware of the need for analytics to support their marketing decisions. We believe our research offers retailers a new approach to utilize the data that is available to them such that they can make better-informed decisions. One of the implications of using our model is the need for retailers to collect and manage longitudinal data. We argue that consumer behavior changes over time; new habits form and these changes differ for each person both in terms of channel usage and the rate of change. Therefore, especially omnichannel retailers need to collect relevant individual level transaction data including channel preferences, responses to marketing activities, social media usage, and any other type of data that could be used in models of purchasing behavior. More importantly, retailers will need to augment their existing marketing teams to include individuals with advanced analytical skills or statisticians to capitalize on their data.

Alternatively, retailers might need to rely more heavily on outsourcing their marketing research activities to consulting agencies that have statistical expertise with *Big Data*. The computational complexities increase exponentially as we develop more sophisticated models of consumer behavior. Market leaders will be those that not only capture the data but also those who invest in continuous analytics to identify change in behaviors and who adapt their marketing activities accordingly.

Another implication of our research is that as new and innovative marketing tactics are introduced into the marketplace, retailers need to be cautious in how they manage their marketing activities. Despite the high costs of direct mailings, we argue this marketing activity continues to be an important activity that should remain in a marketer's toolkit. As new and potentially less costly marketing promotions are developed, for example electronic marketing communications, we argue that retailers will require continuous analytics to understand the relative impact of each marketing activity on managerially relevant outcomes such as customer profitability, purchase incidence, and frequency of purchases. We hope our dynamic segmentation framework offers a framework that can easily be extended to integrate other marketing activities into future research of purchasing behavior.

Limitations and Directions for Future Research

Our research has several limitations. First, at the time of the study, the retailer did not have an Internet marketing or mobile strategy and only tracked catalog spend. Future research would benefit from a model that incorporates other marketing activities so as to understand the relative effectiveness of each promotion. Would coupons generate the same behavior in customers (i.e, migrate to another channel) at a lower cost to the retailer? Is it the marketing activity that triggers a response or is it the availability of more channels that ultimately make a customer more profitable?

Although our research provides insight into the profitability of different types of customer-firm relationships, future research could extend our research by developing an optimization model where the objective is to maximize the value of the customer base. This objective could be achieved by (1) migrating unprofitable customers to channels that are less costly to serve, and (2) using marketing communications and promotions to encourage customers to buy more from a retailer (i.e., use of alternative channels to buy more from the preferred retailer).

Conclusion

In this research, we developed a dynamic segmentation framework to assess purchasing behavior in a multichannel environment. Our framework predicts a customer's purchasing behavior conditional on channel choice while simultaneously accounting for customer heterogeneity, the impact of direct mail, and the evolution of different types of customer-firm relationships. Overall, our research extends our understanding of the differential impact of direct mail on dynamic purchasing behaviors beyond a customer's first purchase. Finally, our findings provide new insight for retailers managing customers in an ominichannel environment.

References

Abhishek, Vibhanshu, Peter Fader, and Kartik Hosanagar (2012), "The Long Road to Online Conversion: A Model of Multichannel Attribution," Available at SSRN 2158421.

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

Ascarza, Eva, and Bruce Hardie (2011), "A Joint Model of Usage and Churn in Contractual Settings," Available at SSRN 1898552.

Avery, Jill, Thomas J. Steenburgh, John Deighton, and Mary Caravella (2012), "Adding Bricks To Clicks: Predicting The Patterns Of Cross-Channel Elasticities Over Time," *Journal of Marketing*, 76 (May), 96-111.

Baksh, Fasil M., Dankmar Böhning and Rattana Lerdsuwansri (2011), "An Extension of an Over-Dispersion Test for Count Data," *Computational Statistics & Data Analysis*, 55, 466–474.

Böhning, Dankmar (1994), "A Note on a Test for Poisson Overdispersion," *Biometrika*, 81, 418–419.

Bulla, Jan, and Andreas Berzel (2008), "Computational Issues In Parameter Estimation For Stationary Hidden Markov Models," *Computational Statistics,* 23(1), 1-18.

Campbell, Dennis, and Frances Frei (2010), "Cost Structure, Customer Profitability, And Retention Implications Of Self-Service Distribution Channels: Evidence From Customer Behavior In An Online Banking Channel," *Management Science*, 56(1), 4-24.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin (1977), "Maximum Likelihood From Incomplete Data Via The EM Algorithm," *Journal of the Royal Statistical Society*. Series B (Methodological), 1-38.

Dinner, Isaac M., Harald J. van Heerde, and Scott A. Neslin (2011), "Driving Online and Offline Sales: The Cross-Channel Effects of Digital Versus Traditional Advertising," Available at SSRN 1955653.

Keller, Kevin Lane, and Donald R. Lehmann (2006), "Brands And Branding: Research Findings And Future Priorities," *Marketing Science*, 25(6), 740-759.

Konuş, Umut, Peter C. Verhoef, and Scott A. Neslin (2008), "Multichannel Shopper Segments And Their Covariates," *Journal of Retailing*, 84(4), 398-413.

Kushwaha, Tarun, and Shankar Venkatesh (2013),"Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics," *Journal of Marketing*, 1-19.

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.

Lyytinen, Kalle and Gregory Rose (2003), "The Disruptive Nature of Information Technology Innovations: The Case of Internet Computing in Systems Development Organisations." *MIS Quarterly*, 27(4), 557-595.

Mark, Tanya, Katherine N. Lemon, Mark Vandenbosch, Jan Bulla and Antonello Maruotti (2013), "Capturing the Evolution of Customer–Firm Relationships: How Customers Become More (or Less) Valuable Over Time." *Journal of Retailing*, 89(3), 231-245.

Min, Yongyi, and Alan Agresti (2005), "Random Effect Models For Repeated Measures Of Zero-Inflated Count Data," *Statistical Modelling,* 5(1), 1-19.

Montoya, Mitzi M., Anne P. Massey, and Vijay Khatri (2010), "Connecting IT Services Operations To Services Marketing Practices," *Journal of Management Information Systems* 26(4), 65-85.

Naik, Prasad A., and Kay Peters (2009), "A Hierarchical Marketing Communications Model Of Online And Offline Media Synergies," *Journal of Interactive Marketing*, 23(4), 288-299.

Narasimhan, Chakravarthi, Scott A. Neslin, and Subrata K. Sen, (1996), “Promotional Elasticities And Category Characteristics,” *Journal of Marketing*, 60(2), 17-30.

Neslin, Scott A., Gail Ayala Taylor, Kimberly D. Grantham, Kimberly R. Mcneil (2013), “Overcoming ‘The Recency Trap’ In Customer Relationship Management," *Journal of the Academy of Marketing Science*, 41(3), 320-337.

_____, 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.

_____, Scott A., Dhruv Grewal, Robert Leghorn, Shankar Venkatesh, et al. (2006), "Challenges And Opportunities In Multichannel Customer Management," *Journal of Service Research*, 9(2), 95-112.

Netzer, Oded, James M. Lattin, and V. Srinivasan, "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*, 27(2), 185-204.

Simester, Duncan I., Peng Sun, and John N. Tsitsiklis (2006), "Dynamic Catalog Mailing Policies," *Management Science*, 52(5), 683-696.

Todem, D., Y. Zhang, A. Ismail, and W. Sohn (2010), "Random Effects Regression Models For Count Data With Excess Zeros In Caries Research," *Journal of Applied Statistics,* 37(10), 1661-1679.

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

Venkatesan, Rajkumar, V. Kumar, and Nalini Ravishanker (2007), "Multichannel Shopping: Causes and Consequences," *Journal of Marketing*, 71(2), 114-132.

Venkatesh,Viswanath, Michael Morris, Gordon B. Davis, Fred D. Davis (2003), "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly,* 27(3), 425-478.

Verhoef, Peter C., Scott A. Neslin, and Bjorn Vroomen (2007), "Multichannel Customer Management: Understanding the Research-Shopper Phenomenon," *International Journal of Research in Marketing*, 24(2), 129-148.

Wiesel, Thorsten, Koen Pauwels, and Joep Arts (2011), "Practice Prize Paper—Marketing's Profit Impact: Quantifying Online and Off-line Funnel Progression," *Marketing Science*, 30(4), 604-611.

Zucchini, Walter, and Iain L. MacDonald (2009), Hidden Markov Models For Time Series:Introduction Using R., CRC Press.

TABLE 1
Summary of Research in Multichannel Literature

Research	*Key Findings*	*Type of Data*	*Length of Observation Period*	*Modeling of Dynamics*	*Number of Channels*	*Marketing Covariates*
Kushwaha and Shankar (2013)	Multichannel customers are more valuable. The moderating impact of product category on strength of the relationship and channel preference.	Cross-sectional	4 years	Utility framework	web and catalog	No
Valentini, Montaguti, and Neslin (2011)	Channel preference is best modeled with two processes: pre and post trial.	Transactional, subscription business model	4.5 years	Two separate logit models one for each stage (pre and post trial) based on learning.	Web, catalog and store	Yes
Ansari, Mela, Neslin (2008)	Model customer channel migration. Web customers are more valuable.	Transactional	4 years	Utility framework	Web and catalog	Yes
Konus, Verhoef, and Neslin (2008)	Emphasize the importance of segmenting customers using channel orientation as a basis.	Survey data	NA	Latent class model Segment based on customers channel preferences for search and purchase activities.	Web, catalog and store	No
Avery et al. (2012)	Assess the effect of introducing new retail store channel on sales. They find, in the long run, sales increase in both web and catalog channels.	Sales data	2 years	Quasi-experimental design	Web, catalog, store	No
Current Research	Multi-channel customers are more valuable and are more likely to respond to marketing activities but it varies by customer-firm relationship type.	Transactional	9 years	Hidden Markov Model	Web, catalog, store	Yes

TABLE 2
Model Selection

	Simple Channel Hurdle	Dynamic Channel Hurdle Model (HMM)				Latent Class Channel Model			
Number of States	-	2	3	4	5	2	3	4	5
Log-likelihood	-88,425	-80,149	-77,512	-75,323	-75,030	-81,727	-79,886	-78,319	-77,273
Number of parameters	48	99	152	207	264	97	146	195	244
AIC	176,947	160,497	155,328	151,060	150,589	163,648	160,064	157,029	155,034
BIC	177,069	161,337	156,618	**152,818**	152,830	164,472	161,303	158,685	157,106

TABLE 3
Hidden Markov Model Parameters

	Transition Matrix				Initial Probabilities	
	State 1	State 2	State 3	State 4		
State 1	***.9522***	***.0176***	***.0093***	***.0209***	State 1	***.8434***
	(.0002)	(.0000)	(.0000)	(.0002)		(.0001)
State 2	***.0060***	***.9567***	***.0302***	***.0071***	State 2	***.1369***
	(.0000)	(.0001)	(.0000)	(.0000)		(.0001)
State 3	***.0009***	***.0063***	***.9926***	***.0002***	State 3	***.0014***
	(.0000)	(.0000)	(.0000)	(.0000)		(.0000)
State 4	***.0118***	***.0160***	***0.0013***	***0.9709***	State 4	***.0183***
	(.0001)	(.0000)	(.0000)	(.0001)		(.0000)

TABLE 4
Inferred Transitions

State				
	State 1	*State 2*	*State 3*	*State 4*
State 1	15491	274	152	293
State 2	31	7155	216	43
State 3	4	24	5405	0
State 4	36	83	6	5787

TABLE 5
Customer dynamics

Number of Transitions	*Proportion*	*Cumulative Frequency*
0	22.30%	22.30%
1	48.90%	71.20%
2	21.10%	92.30%
3	6.10%	98.40%
4	1.30%	99.70%
5	0.20%	99.90%
6	0.10%	100.00%

TABLE 6[ab]
Multinomial Parameter Estimates

State 1				
Parameter	Channel			
	Store $_1$	Phone $_2$	Web $_3$	Across $_4$
Intercept ( )$_1$	***-5.816***	***-1.825***	***-6.275***	***-5.682***
	(.001)	(.023)	(.001)	(.002)
Married ( )$_2$	***.952***	***.214***	***1.447***	***.406***
	(.001)	(.014)	(.001)	(.002)
Gender ( )$_3$	***-.387***	***-.093***	***.475***	***.002***
	(.000)	(.007)	(.000)	(.001)
Catalogs ( )$_4$	***.076***	***.125***	***.046***	***.167***
	(.014)	(.027)	(.016)	(.056)
Lag of Catalogs ( )$_5$	.025	.021	.026	.034
	(.014)	(.056)	(.016)	(.048)
Holidays ( )$_6$	***.244***	***1.367***	***1.095***	***1.992***
	(.000)	(.003)	(.001)	(.001)

State 2				
Parameter	Channel			
	Store $_1$	Phone $_2$	Web $_3$	Across $_4$
Intercept ( )$_1$	***.622***	***-1.293***	***-3.092***	***-1.397***
	(.006)	(.003)	(.003)	(.007)
Married ( )$_2$	***.410***	***-.011***	***.236***	***.105***
	(.005)	(.002)	(.002)	(.005)
Gender ( )$_3$	***.082***	***-.579***	***.190***	***-.180***
	(.002)	(.001)	(.001)	(.002)
Catalogs ( )$_4$	***-.089***	***.063***	***.108***	***.100***
	(.011)	(.024)	(.031)	(.021)
Lag of Catalogs ( )$_5$	***-.037***	.020	.051	.015
	(.012)	(.023)	(.032)	(.030)
Holidays ( )$_6$	***.957***	***1.537***	***1.328***	***2.212***
	(.001)	(.001)	(.001)	(.001)

State 3				
Parameter	Channel			
	Store $_1$	Phone $_2$	Web $_3$	Across $_4$
Intercept ( )$_1$	***-3.252***	***-5.625***	***-6.302***	***-10.033***
	(.003)	(.002)	(.005)	(.002)
Married ( )$_2$	***-.013***	***.314***	***.019***	***-.077***
	(.002)	(.001)	(.003)	(.002)
Gender ( )$_3$	***-.352***	***.451***	***-.147***	***.099***
	(.001)	(.000)	(.002)	(.001)
Catalogs ( )$_4$	-.048	***.243***	***.400***	***.528***
	(.028)	(.017)	(.017)	(.019)
Lag of Catalogs ( )$_5$	-.027	.009	***.185***	***.219***
	(.027)	(.020)	(.019)	(.022)
Holidays ( )$_6$	***.635***	***.395***	***.761***	***1.440***
	(.000)	(.001)	(.001)	(.001)

State 4				
Parameter	Channel			
	Store $_1$	Phone $_2$	Web $_3$	Across $_4$
Intercept ( )$_1$	***-2.347***	***-1.926***	***-0.543***	***-5.137***
	(.001)	(.002)	(.000)	(.001)
Married ( )$_2$	***1.671***	***-.122***	***-.129***	***.361***
	(.001)	(.002)	(.000)	(.001)
Gender ( )$_3$	***.569***	***-.519***	***-.045***	***-.323***
	(.000)	(.001)	(.000)	(.001)
Catalogs ( )$_4$	***-.072***	***.256***	***-1.085***	***.275***
	(.006)	(.037)	(.000)	(.021)
Lag of Catalogs ( )$_5$	***-.117***	.046	***-.886***	***.058***
	(.005)	(.027)	(.000)	(.021)
Holidays ( )$_6$	***.124***	***1.859***	***-0.022***	***2.189***
	(.000)	(.001)	(.000)	(.001)

a Entries in bold and italic script are significantly different to zero with p = 0.05.
b Standard errors are presented below coefficients in parentheses. Standard errors were determined by non-parametric bootstrap.

TABLE 7$_{ab}$
Truncated Poisson Model

State 1				
Parameter	Channel			
	Store β^1	Phone β^2	Web β^3	Across β^4
Intercept $(\beta)_1$	***-2.214***	***-.870***	***-.628***	***.535***
	(.001)	(.038)	(.003)	(.007)
Married $(\beta)_2$	***-.185***	***.086***	***-.410***	***.047***
	(.001)	(.024)	(.002)	(.005)
Gender $(\beta)_3$	***-.402***	***-.102***	***.033***	***-.091***
	(.000)	(.009)	(.001)	(.002)
Catalogs $(\beta)_4$	***.120***	.054	***-.232***	.025
	(.022)	(.171)	(.036)	(.021)
Lag of Catalogs $(\beta)_5$	.001	.006	.053	.002
	(.019)	(.179)	(.035)	(.026)
Holiday $(\beta)_6$	***.758***	***.622***	***1.303***	***.344***
	(.001)	(.022)	(.000)	(.002)

State 2				
Parameter	Channel			
	Store β^1	Phone β^2	Web β^3	Across β^4
Intercept $(\beta)_1$	***.099***	***-.815***	***-.792***	***1.094***
	(.004)	(.010)	(.003)	(.019)
Married $(\beta)_2$	***.248***	***-.024***	***.039***	-.011
	(.004)	(.006)	(.002)	(.012)
Gender $(\beta)_3$	***.323***	***-.314***	***-.201***	***.022***
	(.003)	(.002)	(.001)	(.005)
Catalogs $(\beta)_4$	-.017	.055	***.028***	.015
	(.020)	(.056)	(.012)	(.066)
Lag of Catalogs $(\beta)_5$	.023	-.008	.021	-.002
	(.016)	(.060)	(.011)	(.070)
Holiday $(\beta)_6$	***.253***	***.862***	***.540***	***.243***
	(.003)	(.003)	(.002)	(.006)

State 3				
Parameter	Channel			
	Store β^1	Phone β^2	Web β^3	Across β^4
Intercept $(\beta)_1$	***.443***	***-1.660***	***-1.650***	***-.988***
	(.003)	(.003)	(.006)	(.004)
Married $(\beta)_2$	***.477***	***.041***	***-.079***	***.009***
	(.002)	(.002)	(0.005)	(.003)
Gender $(\beta)_3$	***.116***	***.784***	***-.128***	***.063***
	(.001)	(.000)	(.005)	(.002)
Catalogs $(\beta)_4$	***-.519***	***.262***	***.123***	***.121***
	(.032)	(.027)	(.013)	(.029)
Lag of Catalogs $(\beta)_5$	.000	***-.293***	***.050***	.054
	(.032)	(.032)	(.016)	(.031)
Holiday $(\beta)_6$	***.478***	***.089***	***.505***	***.281***
	(.001)	(.001)	(.006)	(.003)

State 4				
Parameter	Channel			
	Store β^1	Phone β^2	Web β^3	Across β^4
Intercept $(\beta)_1$	***.238***	***-.254***	***-.019***	***.607***
	(.001)	(.008)	(.000)	(.001)
Married $(\beta)_2$	***-.223***	***.069***	-.011	-.003
	(.001)	(.007)	(.000)	(.036)
Gender $(\beta)_3$	***-.438***	***-.120***	***.001***	***-.142***
	(.000)	(.006)	(.000)	(.001)
Catalogs $(\beta)_4$	***.186***	***.060***	***-.215***	.040
	(.022)	(.028)	(.000)	(.025)
Lag of Catalogs $(\beta)_5$	***-.210***	.023	***-.115***	.021
	(.017)	(.023)	(.000)	(.025)
Holiday $(\beta)_6$	***.683***	***.532***	***.005***	***.393***
	(.000)	(.008)	(.000)	(.002)

a Entries in bold and italic script are significantly different to zero with $p = 0.05$.
b Standard errors are presented below coefficients in parentheses. Standard errors were determined by non-parametric bootstrap.

FIGURE 1
Time Trends of Purchasing Activity and Channel Usage

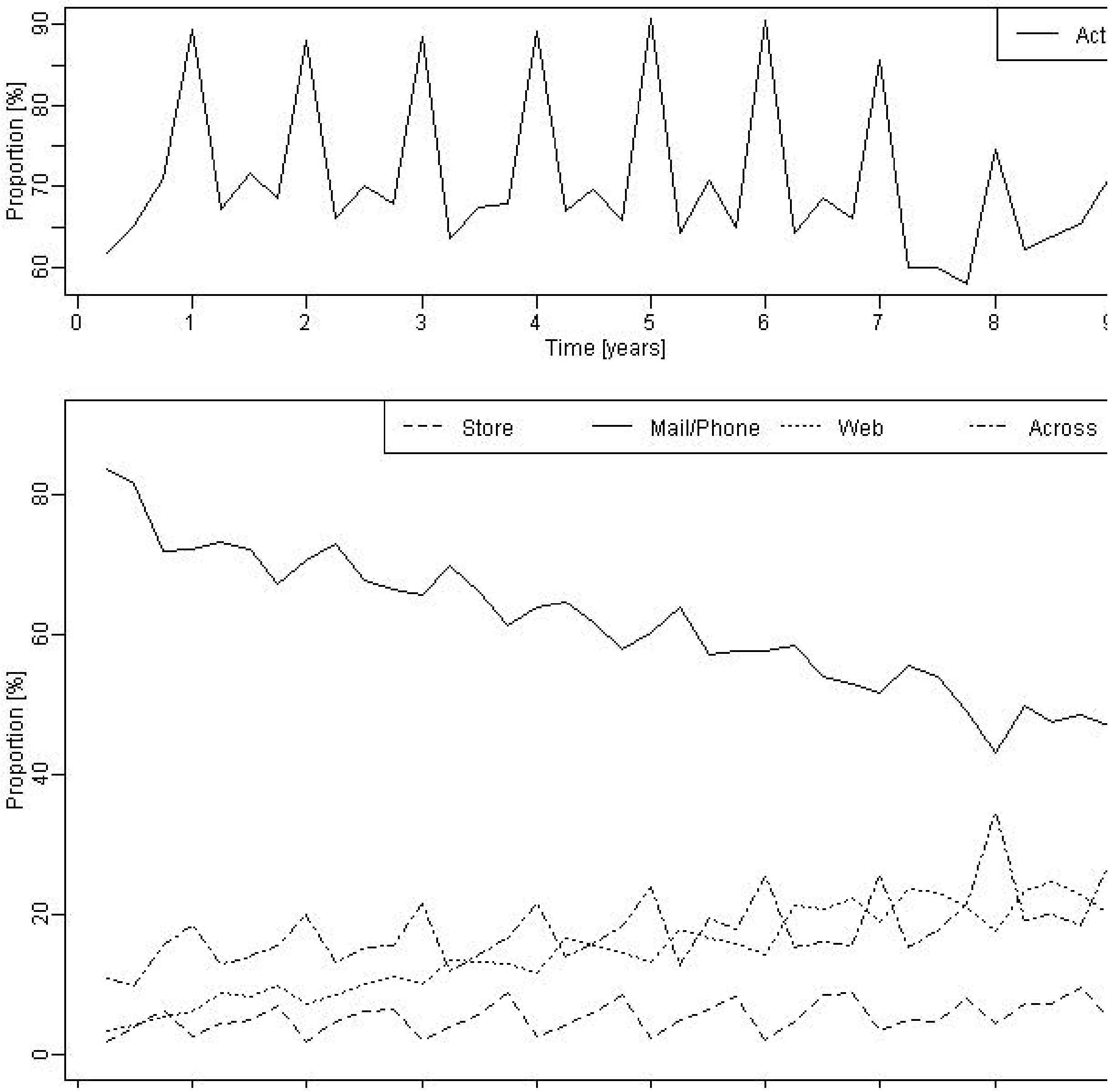

FIGURE 2
Catalog Promotions

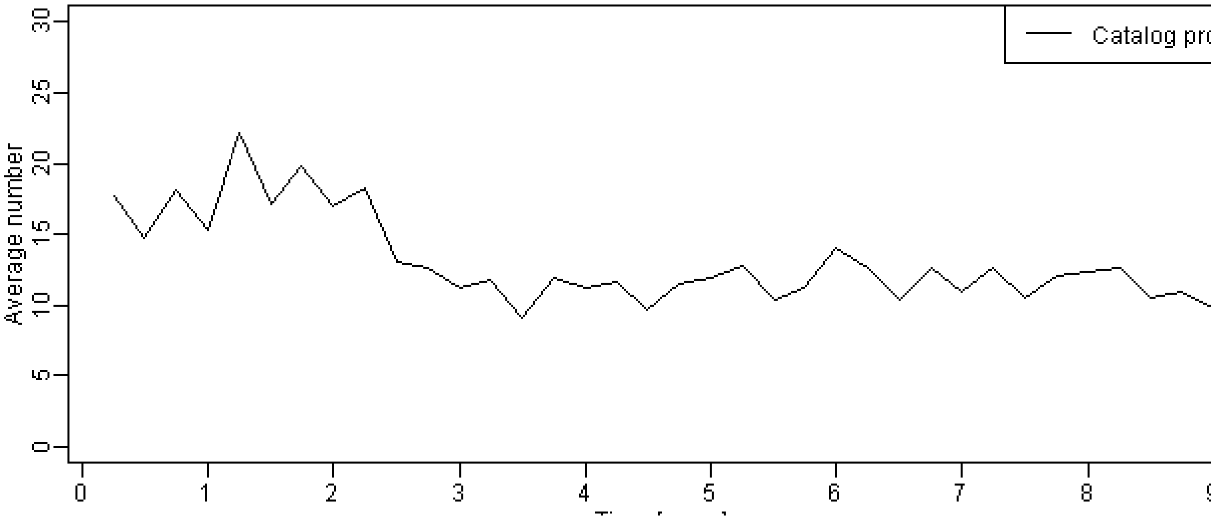

FIGURE 3
Evolution of State Memberships