Customer Evolution in Sales Channel Migration

Sara Valentini
Scott A. Neslin
Elisa Montaguti

*Assistant Professor of Marketing, Management Department, University of Bologna, Via Capo di Lucca, 34 40126 Bologna, Italy, Italy, Phone (+39) 051 209 8400, Fax: (+39) 051 246 411, E-mail: s.valentini@unibo.it

Albert Wesley Frey Professor of Marketing, Tuck School of Business, Dartmouth College, Hanover, NH 03755, USA, Phone: (+1) (603) 646 2841, Fax: (+1) (603) 646 0995, E-mail: scott.a.neslin@dartmouth.edu

Associate Professor of Marketing, Management Department, University of Bologna, Via Capo di Lucca, 34 40126 Bologna, Italy, Phone (+39) 051 209 8089, Fax: (+39) 051 246 411, E-mail: elisa.montaguti@unibo.it
Customer Evolution in Sales Channel Migration

Abstract
Companies often attempt to “right-channel” customers, i.e., encourage them to use the sales channel that adds value both for the firm and the customer. A significant question is when is the best time to right-channel the customer. We develop and estimate a model that sheds light on this issue. The model depicts the evolution of a newly-acquired customer’s channel choice decision process from a “trial” stage to a “post-trial” stage. The model consists of two logit choice models, “trial” and “post-trial”, linked by a geometric process that governs how quickly the customer moves from one to the other. We utilize data for a book retailer who sells through retail stores, the Internet, and catalogs. Our results suggest (1) customers’ decision processes do evolve, (2) within our 4-year window, a minority but sizeable segment switches decision processes, (3) as expected, “switchers” move from a decision process dominated by state dependence and marketing to one driven by channel preference with lower responsiveness to marketing, and (4) the “switcher” segment is multichannel, whereas the “stayer” segment is single channel. (5) identifying these segments can improve right-channeling management.

Keywords: multichannel marketing, channel migration, customer relationship management
INTRODUCTION

The ever-expanding multiplicity of channels through which customers can purchase from companies has made it imperative for managers to understand how customers decide which channels to use (Blattberg, Kim, and Neslin 2008, Chapter 25). One potential benefit of gaining this understanding is the ability to “right-channel” customers, i.e., to induce customers to use the channel or channels that add value both for the customer and the company (Neslin and Shankar 2009). Right-channeling is particularly a challenge when a customer has just been acquired. This may be the time when the customer is malleable and hence receptive to right-channeling efforts, whereas later the customer will become set in his or her ways. Or, acquisition may initiate a “trial” stage whereby the customer naturally will “sample” channels and not be receptive to company marketing efforts until later on. Therefore, an inherent complication in right-channeling is that the customer decision process may evolve – for example, initially be governed by less-entrenched preferences, and evolving toward more preference-based decision-making. The issue is, (1) does the customer’s multichannel decision process evolve after acquisition, (2) if so, in what way, and (3) what are the implications for company right-channeling efforts?

The purpose of this paper is to address these questions. More specifically, we investigate:

- Does the channel choice decision process for a newly acquired customer change over time?
- If so, how predominant is this behavior? I.e., how many new acquired customers are “switchers” and how many are “stayers” with respect to their decision process?
• Among the switchers, how does their initial or “trial” decision process compare with their later or “post-trial” decision process? Do preferences become more dominant? Do customers become more or less receptive to marketing efforts for right-channeling?

• What is the decision process of the stayers and how does it differ from either the trial or post-trial decision process of the switchers?

• What are the implications for company right-channeling policy? Is it easier to right-channel a customer right after acquisition, or difficult at first but become easier over time?

Several papers have investigated the channel choice decision process (e.g., Ansari, Mela, and Neslin 2008; Knox 2006; Thomas and Sullivan 2005; Venkatesan, Kumar, and Ravishanker 2007), reviewed in more detail below. Three key factors that determine channel choice can be distilled from this work: (1) channel preference, (2) company marketing efforts, and (3) channel inertia. The roles that preferences, marketing, and inertia play in channel choice, and their relative dominance, characterizes the customer’s “channel decision process”.

While the above papers have provided key insights on this process, they generally have not considered whether the decision process changes over time. This is particularly relevant in the context of the newly acquired customer, who may for example start out highly inertial but become more preference-based. The notion that the customer’s decision process may evolve is rooted in work such as Heilman et al. 2000 and Meyer and Sathi 1985, and Aaker 1971 (see review below). The relevance of decision-process evolution can be inferred from the trade and managerial press on right-channeling (Myers, Van Metre, and Pickersgill. 2004; Weinberg,
Parise, and Guinan 2007; Shevlin 2007), but our knowledge of decision process evolution and how it should be managed is still rather sparse.

In summary, in order to right-channel newly acquired customers, and in general to know when we should try to right-channel them, we require a model of how the channel choice decision process evolves over time. The objective of this paper is to develop and estimate a model for measuring this evolution, and showing the implications of the model for right-channeling. Our findings indicate that (1) Indeed the customer decision process does change over time, (2) This evolution is exhibited among a significant minority of customers, (3) Among switchers, the decision process changes as we expect, beginning in a relatively inertial but susceptible-to-marketing state, and proceeding to a process more preference-dominated, and (4) switchers tend to be multi-channel customers, whereas stayers tend to be single-channel but this can be changed through marketing activities. Managers can use our framework and results to segment and better target their customers while undertaking right-channeling programs.

The paper proceeds as follows: First we review the literature. This leads to our framework, which serves as a template for categorizing the decision process and charting how it changes over time. We articulate our expectations as to the most likely ways that the process might change. Then we describe the model, the data we use to estimate it, and we present the results. We conclude with implications for researchers and managers.

**LITERATURE REVIEW**

*The Multichannel Decision Process*
Figure 1 shows how researchers have translated the theory by which customers choose channels into quantifiable channel choice models. Blattberg, Kim, and Neslin (2008, Chapter 25) and Neslin et al. (2006) draw on previous research to identify six factors that determine which channel the customer chooses to use: (1) marketing, (2) channel attributes, (3) social influence, (4) channel integration, (5) individual differences, and (6) situational factors.

These factors constitute the “Theory” side of Figure 1. Marketing includes e-mails, catalogs, and other incentives that encourage customers to use one channel versus another. Channel attributes include convenience, privacy, aesthetics, and risk. Social influence entails the impact that significant “others” have on channel choice – e.g., if your friends use the Internet to purchase books, you are more apt to see the merits of the Internet and purchase books there as well. Channel integration refers to coordination among the firm’s channels that makes it easier for the customer to use multiple channels. Individual differences include demographics, psychographics, and previous channel usage experience. Usage experience is particularly important, because it gives rise to channel inertia, the tendency to continue using the same channel used in the past. Channel inertia therefore produces “positive state dependence”. Finally, situational factors are contextual factors present at the purchase occasion, ranging from the weather to the type of purchase (personal vs. gift) to time pressure.

Figure 1 shows how multichannel choice model-builders have organized these factors into marketing, channel preferences, inertia or state dependence, and unobserved factors. Marketing on the left side of Figure 1 maps neatly to marketing on the right side. Channel attributes, channel integration, and social influence map into overall preference. For example,
if Channel A is easy to use, ensures privacy, and is well-integrated with other channels, the
customer is likely to prefer Channel A. Some model-builders have included demographics in
channel choice models, but most importantly, they have utilized various measures of previous
channel usage, i.e., state dependence (e.g., Thomas and Sullivan 2005; Ansari, Mela, and
Neslin 2008). This is often quantified as a lagged dummy variable – e.g., 1 if the customer
used the channel for their last purchase; 0 if not. Finally, situational factors are typically not
observable to the researcher, and hence are not explicitly quantified. Conceptually, therefore,
model-builders have utilized the right-hand side of Figure 1 to estimate channel choice
decision models of the following general form:

\[
\text{Channel Choice} = f(\text{marketing, channel preference, state dependence}) + \epsilon
\]

The determinants of channel choice in equation (1) considered jointly and their relative
dominance (inertia vs. marketing vs. inherent preference) characterize the “channel choice
decision process”.

Several researchers have estimated channel choice models of the form of Equation (1).

Following is a table summarizing these efforts:

<table>
<thead>
<tr>
<th>Research Study</th>
<th>Element of Channel Choice Process</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marketing Channel Preference State Dependence</td>
</tr>
<tr>
<td>Thomas and Sullivan (2005)</td>
<td>✓</td>
</tr>
<tr>
<td>Ansari, Mela, and Neslin (2008)</td>
<td>✓</td>
</tr>
<tr>
<td>Knox (2006)</td>
<td>✓</td>
</tr>
<tr>
<td>Venkatesan, Kumar, and Ravishanker (2007)</td>
<td>✓</td>
</tr>
<tr>
<td>Verhoef, Neslin, and Vrooman (2007)</td>
<td>✓</td>
</tr>
</tbody>
</table>

Thomas and Sullivan, Ansari et al., and Knox, share the common finding that customers are
heterogeneous in the weights they apply to the three elements of the channel choice process.
This is very similar to the brand choice decision models, usually based on supermarket scanner data, which have been pervasive in the literature since the 1980s (see van Heerde and Neslin 2008 for a summary). We will incorporate this heterogeneity in our modelling efforts.

_Evolution in the Channel Choice Decision Process_

With the exception of Knox, none of these papers considers the possibility that the channel choice decision process evolves over time. Knox however focuses especially on the evolution of channel preference and not for example on the evolution of response to marketing efforts. In the literature on brand choice, several papers support the notion that preferences of consumers new to a market vary over time (e.g. Heilman et al. 2000; Meyer and Sathi 1986; Erdem and Keane 1996). For example, Heilman et al. (2000) hypothesized a stage-like process distinguishing among an information collection stage, a middle stage in which information collection continues and is extended to lesser-known brands, and a stage of information consolidation leading to preference for brands that provide the greatest utility. Aaker (1971) pioneered the notion that the decision process could change over time by focusing on the evolution of brand preference. Aaker (1971) modelled preference for a new product in two distinct stages – a “trial” stage, where the consumer is learning about the product, and a “post-trial” stage, where the consumer resolves his or her final preference for the product. Based on Aaker, we consider only two phases of the channel choice process. We believe this is more natural when thinking of the newly acquired customer, but certainly our work could be extended to multi-phase evolution.

The theoretical basis for why the decision process might change over time is fundamentally one of learning. Blattberg, Kim, and Neslin (2008, p. 637; see also Neslin et al. 2006) adapt a traditional decision-making framework to the channel choice context, arguing
that a customer recognizes a channel need, searches for information that addresses that need, and chooses a channel from which to purchase. The customer then evaluates, i.e., learns from, this experience and this shapes how he or she progress through the decision process the next time. The question is why might some customers learn and hence change their decision process. There are at least three reasons:

i. *Motivation and Ability:* Customers vary in their motivation and ability to process information, (Hoch and Deighton 1989; Johnson and Russo 1984; Bettman and Park 1980; MacInnis and Jaworski 1989; Zaichkowsky 1985; MacInnis, Moorman, and Jaworski 1991; Senecal and Nantel 2004). Highly motivated customers want to evaluate different channels and sources of information (e.g. marketing), and those who are able to process this information are likely candidates to switch decision processes. In contrast, unmotivated, uninvolved customers don’t have the motivation, ability, or resources to figure out what might be a more efficient choice process (e.g., see Shugan 1980).

ii. *Familiarity with the Task:* Customers vary in their familiarity with any task (Brucks 1985; Hoch and Deighton 1989; Miyake and Norman 1979; Alba and Hutchinson 1987). For example, certain customers may have already purchased the product category from other companies, and have ample experience with the various channels, with processing marketing communications, etc. They have essentially learned all there is to learn about how to purchase the category, and hence will not change their decision process once they are acquired by another company. Others, however, may be relatively new to the category or new to the particular channels. They have a lot more to learn, and hence will actively circulate through Blattberg, Kim, and Neslin’s feedback loop.

iii. *Unsatisfying Experiences:* Previous research has shown that unsatisfying experiences can have a profound impact on future behavior (Weiner 1986; Mattila 2003). An unsatisfying
experience – a channel was not as convenient as expected, or the information provided in an email was not as reliable as expected, etc. – may cause the customer to re-assess, e.g., try another channel or pay less attention to the next email they receive.

In summary, due to motivation and ability, familiarity with the task, or unsatisfying initial experiences, customers may learn from their initial channel decisions and revise the process by which they make these decisions. I.e., they may change the weights they put on channel preference, marketing, and inertia. One purpose of this paper is to estimate how prevalent it is for newly acquired customers to switch decision processes, and if so, how they change.

**FRAMEWORK**

Our literature review suggests that marketing, channel preference, and state dependence are three key factors of the channel choice decision process, and there is reason to believe that newly acquired customers may change how they weight these three factors, i.e., their decision process may evolve over time. Figure 2 depicts a framework derived from this idea. This framework categorizes customers into one of four possible choice processes, depending on the emphasis they place on marketing, channel preference, and inertia. We see a fundamental difference between preference-based and inertial decision-making, akin to the difference between loyalty and inertia (Engel, Blackwell, and Miniard 1995, p. 158). So our first division is whether customers are “preference-based” (preference weights are higher than state dependence weights) or “inertia-based” (state dependence is weighted higher than channel preference). Within each of those divisions, the consumer might be highly responsive or not highly responsive to marketing. This gives rise to four decision process categories.
Customers in Category 1 have well-established preferences for various channels, but can be influenced by marketing activities to switch channels. Customers in Category 2 have well-established preferences, but cannot be influenced by marketing activities. For example, these people always buy on the Internet and nothing other than unobserved contextual effects (equation 1) can induce them to change.

Customers in Category 3 tend to use the channel they used last time, unless marketing directs them to do otherwise. These customers are habitual but their habits can rather easily be changed. Customers in Category 4 are also inertial, but pay little attention to marketing. They will use the same channel out of habit until an unobserved factor (again, see equation 1) induces them to use a different channel – then they stay with that one. For example, these customers may start off buying from the catalog, but for a particular purchase they may be pressed for time and therefore use the Internet. They then continue using the Internet. Again, this is out of inertia/habit, rather than any true attribute-based loyalty or preference for the Internet.

The interesting part is what switching patterns may occur, i.e., how the channel decision process might evolve. There are 12 possibilities since one can start in one of four categories, and switch to three others. One can make arguments for any one of these patterns to dominate. Eventually, we will measure both how many customers switch processes and which patterns (1 → 2, 3 → 1, etc.) are prevalent in the data. However, our motivation for why we think the decision process might evolve – learning – suggests that customers progress from categories where their preferences are not well formed, i.e., they are predominantly inertial, to
categories where their preferences are well formed and are weighted more importantly than simply what they did last time. Our main expectation therefore is that those who switch their decision processes will generally move from the bottom left of Figure 2 (Categories 3 and 4) to the top left of Figure 2 (Categories 1 and 2).

**THE MODEL**

Our purpose is to model the channel choice decision process and its evolution among a cohort of newly acquired customers. We assume that after being acquired, the customer starts with a *trial* stage multinomial choice model. At some point determined by a geometric distribution, the customer switches to a *post-trial* multinomial choice model. The model therefore consists of two logit choice models linked over time by a geometric process governing when the customer switches from the first to the second. Customers are heterogeneous in their logit model parameters and in the parameter characterizing the geometric distribution. Formally, we call this a *multinomial logit channel selection switching model*.

More specifically, the first component of our model quantifies how long it takes the customer to transit from the *trial* to the *post-trial* logit choice model. We model the length of the *trial* period using a geometric distribution (e.g. Aaker 1971, Buchanan and Morrison 1988; Morrison and Perry 1970; Fourt and Woodlock 1960). The geometric distribution counts the number of trials until an event occurs. It is characterized by a single parameter that signifies the probability the event occurs at each trial. In our case, the “trials” are purchase occasions and the event we are interested in is switching from the *trial* to *post-trial* model. We define the following:
\[ q_h = \] the probability on any given purchase occasion the customer switches from the trial to the post-trial stage choice model, assuming the customer has not switched yet.

\[ X_{ht} = \] the probability the customer has switched to the post-trial stage choice model by the \( t \)-th purchase occasion.

These two quantities are related as follows:

\[
X_{ht} = 1 - (1 - q_h)^{t-1}
\]

This is because \( (1 - q_h)^{t-1} \) is the probability the customer has not switched by the \( t \)-th purchase occasion, so the converse, \( 1 - (1 - q_h)^{t-1} \), is the probability the customer has switched by the \( t \)-th purchase occasion. Note that \( X_{ht} \) is an increasing function of \( q_h \) (the more likely the customer is to switch on any given purchase occasion, the more likely it is that the customer will have switched by purchase occasion \( t \)) and an increasing function of \( t \) (as in all geometric processes, the more trials we have, the higher the likelihood that the event will have occurred).

The second component of our model is the channel choice model. We consider a market with utility-maximizing customers. Customer \( h \) at purchase occasion \( t \) chooses the channel \( j \) (\( j=1,2,\ldots,J \)) among \( J \) channels. \( U_{hjt} \) represents the utility that customer \( h \) derives from choosing channel \( j \) in purchase occasion \( t \). We distinguish between \( U^0 \) and \( U^1 \), where the superscript 0 indicates trial and 1 post-trial. Therefore, \( U^0_{hjt} \) is the utility of choosing channel \( j \) at purchase occasion \( t \), given the customer is in the trial period, and \( U^1_{hjt} \) is the utility of choosing channel \( j \) at purchase occasion \( t \), given the customer is in the post-trial period. These utilities are a function of the elements of the channel choice decision process shown in Equation (1) and Figure 1:

\[
U^0_{hjt} = \alpha_{hj}^0 + \beta_{1hj}^0 CS_{ht} + \beta_{2hj}^0 ES_{ht} + \beta_{3hj}^0 LC_{hjt} + \epsilon_{hjt}^0
\]
where:

\[ \alpha_{hj}^k = \text{Customer } h\text{'s preference for channel } j\text{ during the } tria l\text{ period } (k=0) \text{ or the } post-trial\text{ period } (k=1). \]

\[ CS_{ht} = \text{Number of catalogs sent to customer } h \text{ during the month in which purchase occasion } t\text{ occurs.} \]

\[ ES_{ht} = \text{Number of e-mails sent to customer } h \text{ during the month in which purchase occasion } t\text{ occurs.} \]

\[ LC_{hjt} = \text{1 if customer } h \text{ purchased from channel } j \text{ on the purchase occasion before purchase occasion } t; = 0 \text{ otherwise.} \]

\[ \epsilon_{hjt}^k = \text{Impact of unobserved factors on customer } h\text{'s utility for channel } j\text{ at purchase occasion } t\text{ during either the } tria l\text{ period } (k=0) \text{ or the } post-trial\text{ period } (k=1). \]

\[ \beta_{mhj}^k = \text{Impact of catalogs } (m=1) \text{ or e-mails } (m=2) \text{ on customer } h\text{'s utility for channel } j\text{ during the } tria l\text{ period } (k=0) \text{ or the } post-trial\text{ period } (k=1). \]

\[ \beta_{5h}^k = \text{Impact of previously buying in a particular channel on the utility for that channel, for customer } h \text{ at purchase occasion } t\text{, during the } tria l\text{ period } (k=0) \text{ or the } post-trial\text{ period } (k=1). \]

\[ CS_{ht} \text{ and } ES_{ht} \text{ can depend on the customer\’s past choice, hence can behave like endogenous variables. We used an approach similar to Gönül, Kim, and Shi (2000), motivated by instrumental variables estimation, to minimize endogeneity bias. Specifically, we estimated the probability of receiving a certain number of e-mails with a Poisson regression. We used as explanatory variables seasonality dummies, the lagged number of purchases, and customer characteristics. Similarly, we predicted the number of catalogs that each customer receives. Finally, instead of entering the actual values of } CS_{ht} \text{ and } ES_{ht} \text{ in the multinomial logit channel selection switching model (equations 3 and 4) we used their predicted values} \]
Channel preferences are captured by \( \alpha_{hj}^k \), the importance of marketing is captured by \( \beta_{mhj}^k \), and the importance of inertia or state dependence is captured by \( \beta_{2h}^k \). The relative sizes of these parameters will enable us to classify customers into the categories defined in Figure 2, both before and after they switch decision processes.

The preference and marketing impact parameters are channel specific – customers have different preferences for different channels, and marketing has a differential impact on Channel A versus Channel B (note the marketing variables are not channel specific, as in Thomas and Sullivan 2005). The state dependence parameter is not channel specific, because its independent variable, \( LC_{hjt} \), already varies across channel alternatives. Quantifying state-dependence using a 0-1 dummy variable (\( LC_{hjt} \)) is consistent with the literature (see Thomas and Sullivan 2005). We initialize \( LC_{hjt} \) using the the first purchase occasion and begin the estimation with the second purchase occasion.

Using equation (2) and the form of the multinomial logit model, we can put together the geometric evolution process and the decision processes for the trial and post-trial periods into one unifying equation for the probability the customer chooses a particular channel. Equation 5 shows this multinomial channel choice switching model, which computes the probability that customer \( h \) chooses channel \( j \) during purchase occasion \( t \) (\( PrC_{hjt} \)):

\[
PrC_{hjt} = (1 - X_{ht}) \left( \frac{\exp(U_{hjt}^0)}{\sum_{j=1}^{3} \exp(U_{hjt}^0)} \right) + X_{ht} \left( \frac{\exp(U_{hjt}^1)}{\sum_{j=1}^{3} \exp(U_{hjt}^1)} \right)
\]

Note that on the customer’s first purchase occasion, \( t = 1 \) so \( X_{ht} = 0 \) (equation 5) and the customer’s decision process is driven entirely by the trial model. Over time, \( X_{ht} \) increases,
the degree to which is determined by the customer geometric parameter $q_h$, and the process is
driven more by the post-trial model. For any purchase occasion, the customer’s probability of
choosing channel $j$ is a weighted function of a trial and post-trial probability model, the
weights being the likelihood that the customer is in the trial stage or the post-trial stage,
respectively. Since the weight placed on the post-trial model increases as the customer
accumulates more purchases (equation 2), the post-trial model becomes more important as
purchases accumulate, and the trial model becomes less important. Exactly how fast this
occurs depends not only on the number of purchases accumulated but on the parameter $q_h$.

Equations (2) – (5) capture the phenomenon we want to measure, the customer’s
evolution from one decision process to another. Note we assume the geometric parameter $q_h$
varies across customers but not over time. Recall we identified motivation and ability, task
familiarity, and unsatisfying experiences as the bases for suspecting that customers might
switch decision processes. We consider ability as a customer trait, suggesting $q_h$ would be
constant over time. However, unsatisfying experiences and task familiarity would be time
dependent. Motivation could change as well, for example, as a result of marketing. Hence, it
is possible that $q_h$ could change over time. We test additional models to explore this
possibility. We find these models do not improve on the “simpler” model represented by
equations (2) – (5), suggesting $q_h$ is largely a consumer trait.

ESTIMATION AND DATA

We use hierarchical Bayesian estimation for the multinomial channel choice switching
model. This approach is particularly suited for panel data sets with a small amount of
information per decision unit when compared to the amount of information available from a
cross-section of many decision units. It enables us to account for parameter heterogeneity, and
provides a convenient method to obtain customer-level estimates of the parameters (Rossi and
Allenby 2003). We hypothesize that the "first stage" priors of the considered parameters
follow a normal distribution with mean $\mu$ and precision $\tau$. The "second stage" priors follow a
normal distribution for $\mu$ and a gamma distribution for $\tau$ (i.e. an inverse gamma distribution on
the variance)\textsuperscript{iii}. For identification purposes we set one channel (the store) as the base. Thus,
we will have alternative-specific coefficients only for $J-1$ channels, scaled relative to the store.

We use data from a major European book retailer operating in one country. The retailer
utilizes three sales channels – physical retail stores, catalogs, and the Internet. The company
operates a subscription-oriented business, thus each customer must become a member in order to
purchase. A code is associated with each customer, tracking each time she or he purchases and
from which channel. We therefore have information on: which channel was selected by each
customer during each purchase occasion, the date of each purchase occasion, how much was
spent, the number and types of communications (e-mails and catalogs) each customer received
and the exact time during which customers received them, and age and gender demographic
characteristics. The period of observation is January 2002 - June 2006\textsuperscript{iv}. We sample a cohort of
new customers who live in at least one store’s service area and who entered into a subscription
agreement with the company during the same period – the 4\textsuperscript{th} quarter of 2001. During the
considered period some new physical retail stores opened. A new store opening causes a
modification in the customers’ choice set. While technically our model could handle this as long
as the multinomial logit models did not violate the Independence of Irrelevant Alternative (IIA)
assumption, our interest is in how customers’ decision process changes with respect to a constant
set of channel alternatives. Therefore, we focus our attention on a static sample of active
customers (i.e. customers having an active relationship with the company between January 2002 and June 2006) with a “full channel choice set” available throughout the relationship.

The final sample size is 1018 households. 70.6% of whom are “single-channel”– they mainly use the same channel over time but may use other channels on one or two purchase occasions. Specifically, 27.6% use mainly the catalog, 42.4% use mainly the stores and 0.6% mainly the Internet (see Table 1). 29.4% are multi-channel customers (4.3% use three channels, and 25% two channels). Table 1 presents additional descriptive information about the sample. Figure 3 shows that the proportion of customers receiving e-mails is roughly the same over time.

RESULTS

Model Estimation and Fit

We use 10,000 iterations for estimation burn-in and one million iterations for estimation. We assess convergence for multiple chains using the Brooks-Gelman-Rubin convergence statistic (Brooks and Gelman 1998) and history graphs. Both suggest strong convergence. The Appendix contains the history graphs. We estimated five different versions of the model. Model 1 is a single multinomial logit model that does not distinguish between trial and post-trial stages. Model 2 is a multinomial logit that assumes the customer starts off using one multinomial logit model, and then switches to another multinomial logit model after an a priori-defined number of purchase occasions. Model 3 is the multinomial logit channel selection switching model described in equations 2-5. Model 4 is a variation of the multinomial logit channel selection
switching model, allowing $q_h$, the probability the customer switches decision processes on any given purchase occasion, to be a function of customers’ demographic characteristics (gender and age). Models 5 and 6 are expanded versions of Model 4, where we allow $q_h$ to vary over time. Model 5 assumes that $q_{ht}$ is affected by marketing variables (total number of catalogs and e-mails received by customer $h$ in each period). In Model 6a $q_{ht}$ varies as a function of demographics (age and gender), lagged number of channels used, and lagged product returns (which may indicate unsatisfying experiences and hence motivation to learn about new channels). Model 6b considers just lagged returns as a covariate.

We compare these seven models using the deviance information criterion (DIC) statistic (Spiegelhalter et al. 2002), a measure of model complexity and fit. Table 2 shows the DIC statistic for the seven models. The best is Model 3, which assumes distinct trial and post-trial logit models, and the length of the trial period is heterogeneous among customers but not systematically a function of customer characteristics or marketing efforts. Model 3’s superiority to Model 1 suggests that indeed, two distinct stages (trial and post-trial) characterize channel choice. A model that accounts for the potential for customers to switch decision processes fits better than one that assumes the customer does not switch. Model 3’s superiority to Model 2 suggests that the length of the trial period is heterogeneous among customers. Model 3’s superiority to Model 4 suggests that customers’ demographics do not affect the probability to switch and the length of the trial period. Model 3’s superiority to Model 5 suggests that marketing does not significantly affect the length of the trial period and the probability of switching one’s channel decision process. Finally, while Model 6a and Model 6b have the hypothesized positive sign for lagged returns, these models as a whole did not improve upon
Model 3. In summary, the model fits of alternative models did not improve over Model 3. We therefore use Model 3 as our basis for further analysis.

[Insert Table 2 about here]

Prevalence of Decision Process Switching, and Characteristics of Switchers vs. Stayers

While our comparative model testing suggests that customers switch decision processes, we want to know how prevalent this behavior is, i.e., how many customers switch processes. To obtain this information we substituted the estimated $q_h$ for each customer, together with each customer’s total number of purchases ($T_h$) into equation (2) to compute the probability that each customer switches by her last purchase occasion (i.e., we used $T_h$ for $t$ in equation 2). This was calculated at each iteration of the Bayesian estimation process and then averaged for each customer to obtain the probability the customer would have switched processes by the end of the observation period. Descriptive statistics for this number are displayed in Table 3.

[Insert Table 3 about here]

The first row of Table 3 shows we expect 22% of customers to have switched by their last purchase occasion. To determine which particular customers would be the switchers, we ordered customers according to their probability of switching by the end of the data, and then took the 22% highest as the switchers. The logic here is that the customers with the highest probability of switching are the ones we would expect to switch. The remaining customers were classified as “stayers,” i.e., not expected to switch by the end of the data. The last two rows in Table 3 show the average probability of switching for stayers and switchers. This shows a clear separation between the two groups.
Table 4 examines the usage of various channels by *switchers* versus *stayers*, calculated directly from the data for *switchers* and *stayers*. The table presents an intriguing finding, namely, that *stayers* are mainly single-channel users, while *switchers* are predominantly multi-channel. In retrospect, this is a plausible result based on the theory we discussed early for why customers might switch their decision process: A *switcher* would have the motivation and ability to learn, would be unfamiliar with the task environment, or would have had an unsatisfactory experience. These factors and multi-channel usage would seem to go hand-in-hand. Motivation and ability to learn would encourage experimentation with various channels. Unfamiliarity with channels would also encourage experimentation. And of course, an unsatisfactory experience would encourage the customer to consider alternative channels. In any event, this is an important finding and a characterization of the multi-channel customer previously not demonstrated before—the multi-channel customer is one who switches decision processes over time.

[Insert Table 4 about here]

The estimates for the mean parameters for the multinomial channel choice switching model (Model 3) are presented in Table 5vii.

[Table 5 goes here]

The intercept estimates suggest a preference for the use of the store over the Internet in the *trial* period, a preference that increases in the *post-trial* period. Interestingly, the intercept for the catalog over the store is not significant (i.e. the 95% posterior confidence interval excludes zero), suggesting that on average we cannot distinguish a stronger preference for the use of the catalog versus the store.

Table 5 reports parameter estimates and elasticities for marketing activitiesviii. The effect of catalogs is always significant in the *trial* period. It has a positive association with store choice, suggesting that catalogs sent “promote” the use of the store. This result is interesting because
catalogs sent did not have special promotions that featured the physical store, however they contained a list of the stores with a picture of one “representative” store. This effect is particularly strong in the choice of the store over the Internet. In the post-trial period however, the effect of catalogs is not significant. In the trial period, e-mails do not influence the choice of the catalog over the store. However, they increase the probability of the use of the Internet over the store. In the post-trial period; e-mails slightly affect the choice of the catalog over the store, but no longer increase the probability to choose the Internet over the store. In general, marketing on average becomes somewhat less effective in the post-trial model. This is seen especially for catalogs and somewhat for e-mails.

Finally, it is worth noting that the effect of state dependence is reduced in the post-trial period, suggesting that on average people become less inertial. The increasing of the Internet vs. store effect, combined with the decreasing of the state dependence parameter in the post-trial stage, suggests that on average post-trial channel choice is guided more by channel preferences than by inertia.

In summary, these results show that the parameter estimates differ in important ways from the trial to the post-trial phase. The decision process changes in two main ways: i) channel preferences become more important and state dependence less important over time; and ii) marketing communication generally becomes less effective in the post-trial stage.

**Diagnosing How the Decision Process Changes among Switchers**

We now segment customers in terms of the framework depicted in Figure 2, i.e., quantify the extent that different decision processes are employed in the trial vs. post-trial stages, and examine the patterns by which switchers migrate from one process to another.
The Bayesian estimation yields individual parameter estimates for the impact of marketing communications, for channel preferences and for the impact of state dependence. This allows us to classify customers into the channel decision style categories in Figure 2: categories 1) preference-based decision-making coupled with high marketing responsiveness, 2) preference-based decision-making and low marketing responsiveness, 3) inertial decision-making and high marketing responsiveness, 4) inertial decision-making and low marketing responsiveness.

By comparing the magnitudes of the channel preference parameters with the state dependence parameters, we can classify customers as belonging to either preference-based or inertia-based decision strategy’s groups. Specifically, we use the following rule to classify customers: if customer $h$ exhibits a positive and high state dependence – higher than her preference for the catalog over the store and the Internet over the store (in absolute values) – we classify this customer as inertial (see Table 6 for details).

After having classified customers as preference-based or inertia-based, we further distinguish them on the basis of their marketing responsiveness. In order to do so, we consider marketing elasticities. We compare individual marketing elasticities with the median value across customers, and we identify customers as belonging to the high marketing responsiveness group if e-mails or catalogs sent elasticities are greater than the respective median elasticities (see Table 7 for details).

We used the above decision rules to classify each switcher based on her or his parameters and elasticities into one of the four decision process categories. We used customers’ estimated trial parameters to classify them for the trial stage, and their post-trial parameters to classify them for the post-trial stage. We then “cross-tabbed” category membership by decision stage to
examine the patterns by which customers migrate from one decision process category to another. Figure 4 shows these migrations.

[Insert Figure 4 about here]

The key finding is as expected, a tendency to migrate from inertia-based to preference-based decision-making, and some tendency to migrate from high marketing responsiveness to lower marketing responsiveness. In particular, Figure 4 highlights two main types of migrations: the migration from trial category 3 to post-trial category 1 (38% of all switchers); and the migration from trial category 3 to post-trial category 2 (55%). Both these migrations are from an inertia-based decision strategy to a preference-based decision process.

Customers in category 3 are responsive to marketing as well as inertial. This suggests that marketing bounces them from channel to channel during the trial stage. However, these customers eventually form real preferences for the various channels. This makes them preference-based decision-makers. Some of them remain responsive to marketing, i.e., marketing can still influence their choice, but for the majority, marketing diminishes in importance.

Table 8 shows the distribution of channel usage for the switchers before and after they are expected to switch. The table also shows the average trial and post-trial order size. Interestingly, this table shows that switchers were always multi-channel customers, in other words they do not start out as single channel users and then become multi-channel. Following our interpretation of Figure 4, while switchers begin and end multi-channel, they are multi-channel for different reasons. In the trial stage, they are buffeted among channels mainly by marketing – they inertially stick with the channel they used last time unless acted upon by an outside force, namely marketing. In the post-trial stage, marketing still may have an impact, but the switchers have formed more definite preferences for the channels they like (relative to the impact of state dependence), and utilize the channels for which they have high preference. It turns out that the
pattern of preferences makes them “multi-channel” loyal, similar to “multi-brand” loyalty that can occur in traditional brand choice models if consumers have high brand intercepts for more than one brand.

Policy Simulation to Investigate Right-Channeling Strategies

We have found two segments: the switchers, who start off largely inertial but marketing sensitive, and evolve to a preference-based decision process where marketing is less influential, and the stayers, who are predominantly inertial but influenced by marketing. In this section, we consider the case that the firm wants to “right-channel” all its customers to the Internet using an e-mail campaign. The issue is how effectively they can do this, and what impact the timing of the email campaign has, especially on the switcher segment.

We investigate two e-mail right-channeling strategies by conducting a policy simulation. The first strategy – “heavy-early” – weights emailing toward the purchase occasions immediately after the customer has been acquired. The second strategy – “heavy-late” – weights emailing toward later purchase occasions. Given our earlier results, we expect the heavy-early strategy to be effective, while heavy-late will be ineffective, in inducing switchers to use the Internet. Both strategies should be equally effective with the Stayers, because their decision process does not change. We use the parameter estimates obtained at the individual level and simulate the channel choices that customers make. Figure 5 summarizes the results.

Figure 5A shows that emails play a role in right-channeling switchers, but it is crucial that the campaign be implemented soon after the customer is acquired. The left side of Figure 5A shows that newly acquired can be moved effectively to the Internet using the heavy-early
strategy. However, under the heavy-late strategy, the right side of Figure 5A shows that the switchers will not be induced to choose the Internet. Again, this follows directly from our results – the switchers start off inertial but receptive to marketing, but end up preference-based and less receptive to marketing.

Figure 5B shows that stayers can be induced to use the Internet whether emails are sent immediately after customer acquisition or later. This is because stayers are inertial but marketing sensitive, and stay that way.

Figure 5 shows that companies can right-channel customers, but with the switching segment, one has to be careful about the timing. Given our model findings, the time to right-channel the switchers to the Internet is soon after they are acquired, before their decision process evolves to a preference-based, less marketing responsive state. The clear results in Figure 5A clearly show the value of understanding and measuring the evolution of the channel choice decision process.

CONCLUSIONS

Key Results and Implications

The central questions of this paper were: do newly acquired customers switch to a new channel choice process over time? How prevalent is such change? How can companies use the insights from the answers to these questions to right-channel customers?

We first developed a framework for characterizing “channel choice process”. We employed three main factors that had been examined by previous researchers – channel preference, state dependence or inertia, and marketing – to demarcate four basic processes: preference-based marketing responsive, preference-based not marketing responsive, inertia-based
marketing responsive, and inertia-based not marketing responsive. We hypothesized that due to customers’ motivation and ability to learn, their familiar with the task environment, or unsatisfactory experiences, at least a segment of customers would evolve from one decision process to another.

We developed a multinomial channel selection switching model to test whether this occurs and to answer our key questions. The model posited two multinomial logit choice models, where the customer could evolve from one model to another according to a geometric process. We allowed the parameter that characterizes the geometric process – the probability the customer switches at any purchase occasion – to vary across customers. This approach allowed us to estimate two sets of parameters, one for the trial stage and the other for the post-trial stage. We used Bayesian estimation to derive individual-level parameters, and we used these estimates to classify customers into our decision process taxonomy, and to map migrations between categories over time.

We applied the model to a contractual setting in the book retail industry in which customers’ channel choices were observed over time. Our key findings are:

- Newly acquired customers do switch decision strategies over time.
- Within a moderate length of time (the four years of our observation period), 22% of customers would be expected to switch processes. Therefore it is a segment of customers that actually switches within a given time period.
- The switcher segment consists of multi-channel customers, while the stayer segment tends to be single-channel.
• The predominant pattern of evolution is from an inertia-based, marketing responsive process to a preference-based process that has both marketing responsive and non-marketing responsive customers, but on average, less marketing responsiveness.

• To right-channel customers, it is important to work on the *switchers* soon after they have been acquired. The *stayers* can also be right-channeled, but the timing of the marketing efforts is not crucial. These particular results are idiosyncratic to our data and application, but they point out the importance of quantifying the nature and degree of decision process evolution.

Probably the most important modeling implication of this work is that the newly acquired customer’s channel choice process evolves over time, and therefore should be considered in models of customer channel choice. We do not advocate our model as the *only* way to model this evolution, but the main point is that evolution should be included in some way. We believe the taxonomy we developed for characterizing the channel choice process (Figure 2) is useful and produced a revealing portrait of choice process evolution, namely that customers start off as creatures of habit but can be influenced by marketing, and gradually evolve to a decision process where their preferences are firm and they are less susceptible to marketing as a determinant of which channel they use.

The finding that multi-channel usage and process evolution go together, whereas single-channel usage and process constancy go together, sheds new light on the multi-channel customer. There have been several characterizations of the multi-channel customer (e.g., see Kumar and Venkatesan 2005). The additional insight that the multi-channel customer’s decision process is more likely to be in a state of flux adds to our understanding of that customer.

While some previous work, such as Knox (2006) has considered the possibility that newly acquired customers evolve in their decision process, we examine the possibility that some customers evolve while others do not. From a modeling perspective, the key point is that not
only should decision process evolution be considered, but the tendency to evolve is heterogeneous across customers.

In addition to heterogeneity in the tendency to be a switcher or a stayer, not all switchers are alike. While most end up preference-based; some are marketing responsive while others are not. And even within these two groups, there is variation in marketing responsiveness as well as the particular channels preferred.

Another important result concerns the nature of the marketing communications’ effect. Both Ansari, Mela, and Neslin (2008) and Knox (2006) find that e-mails are strongly associated with choice of the Internet. Similarly Pauwels and Neslin (2008) document that e-mails equally affect catalog and Internet sales, while having marginal impact on store sales. We contribute to this literature by showing that marketing’s effect varies in the trial and post-trial stages.

Managerially our main contribution refers to customer channel management based on the customer channel choice decision process. Specifically, our framework helps gain insights on how customers can be right-channeled through the right level of marketing at the right time. In particular, firms need to: i) distinguish between stayers and switchers, ii) quantify how sensitivity to marketing differs over time for the switchers, and iii) quantify the stayers’ decision process, in particular, how marketing influences their choice of channel.

In general, newly acquired customers who are switchers become more set in their ways and less reliant on marketing. Therefore, the time for firms to right-channel these customers is right after they become a customer. Our model helps managers predict the duration of the trial phase, which defines the length of the high marketing responsiveness period. This might allow managers to specify varying marketing communications time horizons across consumers. Stayers stick with their initial decision process, although we find they tend to be single-channel shoppers.
But they are marketing sensitive and so firms can use marketing to modify their use of various channels.

Limitations and Future Research

Our data refer to one major retailer that operates in a specific industry. Therefore, while our results are intuitive and largely in keeping with our expectations, further work is needed to generalize our empirical findings – in particular, (1) the size of the switcher segment, (2) the nature of switcher evolution from inertial/marketing sensitive to preference-based, less marketing sensitive, and (3) that the switchers are inherently multichannel, while the stayers are inherently single channel.

We used catalogs sent and e-mails sent as independent variables in our model; however, they are very gross measures, given that firms send so many different kinds of catalogs and e-mails. Future research should account for different types of catalogs and e-mails.

In addition, we use data on a subscription-oriented business model that might present distinctive characteristics. The advantages of using this kind of data set are several (for example, we are sure to track all the purchases ever made by the customers in all channels). However, further research is needed to investigate the evolution phenomenon, by distinguishing between the subscription-oriented business model and other types of business strategies.

Finally, we sampled only customers with a “full cycle” to perform our analysis, this to guarantee that customers were not “active” only during the trial stage. In so doing we might have sampled the “best customers”. This limitation might lead to some possible extension of this work. Specifically, we could jointly model customer retention, choice process evolution, and profitability.
REFERENCES


Miyake, Naomi and Donald A. Norman (1979), “To Ask a Question, One Must Know Enough to Know What is Not Known,” *Journal of Verbal Learning and Verbal Behavior*, 18, 357-364.


FOOTNOTES

\footnote{Note that the figure also shows the possibility that the process does not switch; that is, the customer is a “stayer.” This is signified by the solid lines in Figure 2.}

\footnote{We also estimated equation 5 with interactions effect and/or diminishing returns to marketing investments (e.g. Thomas and Sullivan 2005, Venkatesan, Kumar, and Ravishanker 2007). We also tested a model with a catalog stock variable that essentially creates wearout as the customer gets more catalogs. However, the fit of the model was lower and the interpretation of the parameters estimates was essentially identical.}

\footnote{We use a vague prior for the prior mean, e.g. \( \mu \sim \text{Normal}(0,0.00001) \), and a proper but vague inverse-gamma prior on the variance, e.g. \( \tau \sim \Gamma(0.5,0.5) \).}

\footnote{We aggregate data to the quarterly level since the quarterly sampling rate corresponds largely to the decision processes we model. When using quarterly aggregation, multiple purchases are negligible. When there are multiple purchases in the same quarter, we classify the channel with the higher order-size as the channel of choice.}

\footnote{23.5\% were customers using a portfolio of channels over the observation period whereas 5.9\% are individuals who switched channel just once during the observation period (e.g. they were loyal to the catalog and then moved to the Internet).}

\footnote{Specifically, we set the \textit{a priori} period to be nine purchase occasions.}

\footnote{We verified that the estimation procedure was capable of recovering known model parameters through simulation assuming the mean parameters in Table 5 were the true parameters. We simulated different datasets by drawing individual parameters from a normal distribution with the means shown in Table 5. We generated the datasets using (1) randomly generated catalogs sent and e-mails sent variables, or (2) the catalogs and e-mails observed in the data. Using the first dataset the model recovered the true parameters very well. The estimated coefficients were \( \alpha = 0.62, \beta_1 = 1.25, \alpha_2 = -1.91, \beta_1 = -2.22, \beta_2 = -1.57, \beta_3 = -0.23, \beta_4 = -7.55, \beta_5 = -0.55 \).}

\footnote{Indicates that the parameter is significant at 5\%). The correlation between the estimated coefficients in Table 5 and those recovered was 0.98. The coefficients that were statistically significant in Table 5 were also significant in the simulation and with the right sign. The model recovered the true parameters fairly well using the second dataset – the correlation between the coefficients in Table 5 and those recovered was 0.72. We believe that multicollinearity between catalogs and emails explains the lower, although still highly positive, correlation.}

\footnote{Specifically, we compute direct marketing communication elasticities as follows (see Guadagni and Little 1983, appendix 2): \( c_k = b \times X_k \times (1 - m_k) \), where \( X \) represents the elasticity, \( b \) is the coefficient of direct marketing communication, \( X \) represents the average direct marketing communication sent and \( m \) represents the expected share of channel \( k \). We compute elasticity at individual level; hence, we assess the impact of a change in direct marketing communication on each customer’s response outcome. In Table 5, we report the average elasticities.}

\footnote{A customer is classified as higher marketing responsive if any of their four marketing elasticities (catalogs: catalog vs. store, catalogs: Internet vs. store; e-mails: catalog vs. store; e-mails: Internet vs. store) is above the median for that elasticity across customers. As long as the customer is exceptionally responsive to some form of marketing, we classify him or her as highly responsive. While we could have used other classification rules, the important issue is not the level of responsiveness per se, but rather how marketing responsiveness changes from the trial to the post-trial stage. Using this rule we find that 100\% of the switchers are highly marketing responsive in the trial period and 44\% in the post-trial period. We also tested a more restrictive classification rule, requiring two marketing elasticities, not one, to be greater than the median. Obviously, using this rule the number of responsive consumers decreases (from 100\% to 91.6\% in the trial period, and from 44\% to 3.6\% post-trial), but the same general finding holds, that switchers are less marketing sensitive after they switch.}

\footnote{These findings are based on a simulation of the switchers through their decision model, including the logit models to simulate their channel choices, and the value of \( q_h \) to simulate when they switch. We needed to use simulation here because the data do not directly tell us when a customer switches decision processes; this must be inferred from the customer’s estimated parameters.}
Table 1

Descriptive Statistics of the Selected Cohort of Customers (n=1018)

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number of purchase occasions (per year)</td>
<td>3.1</td>
<td>1.1</td>
<td>0.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Average number of purchase occasions over relationship</td>
<td>14.7</td>
<td>2.9</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Average $ returns (per quarter)</td>
<td>0.4</td>
<td>3.5</td>
<td>0</td>
<td>104.4</td>
</tr>
<tr>
<td>Average number of catalogs received (per quarter)</td>
<td>2.0</td>
<td>0.7</td>
<td>0.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Average number of e-mails received (per quarter)</td>
<td>0.7</td>
<td>2.0</td>
<td>0.0</td>
<td>12.0</td>
</tr>
<tr>
<td>Age (Years)</td>
<td>43.8</td>
<td>14.6</td>
<td>21.0</td>
<td>88.0</td>
</tr>
<tr>
<td>Gender (Male)</td>
<td>35.2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Channel Usage*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mainly Catalog</td>
<td>27.6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mainly Internet</td>
<td>0.6%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mainly Store</td>
<td>42.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catalog and Store</td>
<td>15.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catalog and Internet</td>
<td>8.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet and Store</td>
<td>1.2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catalog, Internet, and Store</td>
<td>4.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* “Mainly” means that at least 95% of purchases were made on that channel. Similarly, we classified as two channel users/three channel users those customers who made at least 95% of purchases using two channels/three channels.
### Table 2

**Model Fit Comparison**

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Simple multinomial logit</td>
<td>6881.0</td>
</tr>
<tr>
<td>Model 2</td>
<td>Two multinomial logit models with switching timing pre-set for purchase occasion 9.</td>
<td>6648.2</td>
</tr>
<tr>
<td>Model 3</td>
<td>Multinomial logit channel selection switching model (Equations 2-5)</td>
<td><strong>6525.8</strong></td>
</tr>
<tr>
<td>Model 4</td>
<td>Multinomial logit channel selection switching model (with $q_h$ a function of age and gender)</td>
<td>6598.2</td>
</tr>
<tr>
<td>Model 5</td>
<td>Multinomial logit channel selection switching model (with $q_{ht}$ a function marketing)</td>
<td>6999.4</td>
</tr>
<tr>
<td>Model 6a</td>
<td>Multinomial logit channel selection switching model (with $q_{ht}$ a function of age, gender, lagged number of channel used, and lagged returns)</td>
<td>7284.8</td>
</tr>
<tr>
<td>Model 6b</td>
<td>Multinomial logit channel selection switching model (with $q_{ht}$ a function of lagged returns)</td>
<td>7202.8</td>
</tr>
</tbody>
</table>
Table 3

Summary Statistics: Probability of Switching to the Post-Trial Stage

<table>
<thead>
<tr>
<th>Sample</th>
<th>Freq. (%)</th>
<th>M</th>
<th>Med</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>1018 (100%)</td>
<td>0.22</td>
<td>0.17</td>
<td>0.17</td>
<td>0.08</td>
<td>0.94</td>
</tr>
<tr>
<td><strong>Stayers</strong></td>
<td>793 (78%)</td>
<td>0.14</td>
<td>0.14</td>
<td>0.04</td>
<td>0.08</td>
<td>0.23</td>
</tr>
<tr>
<td><strong>Switchers</strong></td>
<td>225 (22%)</td>
<td>0.49</td>
<td>0.45</td>
<td>0.19</td>
<td>0.23</td>
<td>0.94</td>
</tr>
</tbody>
</table>

*This is calculated using equation (2), using the value of $q_h$ estimated for each customer, and the number of purchases for each customer ($T_h$).*
Table 4

Stayers versus Switchers: Percentage of Times Each Channel is Chosen*

<table>
<thead>
<tr>
<th>Channel Usage</th>
<th>Stayers (793 customers)</th>
<th>Switchers (225 customers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainly Catalog</td>
<td>35.4%</td>
<td>-</td>
</tr>
<tr>
<td>Mainly Internet</td>
<td>0.8%</td>
<td>-</td>
</tr>
<tr>
<td>Mainly Store</td>
<td>54.5%</td>
<td>-</td>
</tr>
<tr>
<td>Catalog and Store</td>
<td>6.2%</td>
<td>48.0%</td>
</tr>
<tr>
<td>Catalog and Internet</td>
<td>2.6%</td>
<td>28.9%</td>
</tr>
<tr>
<td>Internet and Store</td>
<td>0.1%</td>
<td>4.9%</td>
</tr>
<tr>
<td>Catalog, Internet and Store</td>
<td>0.4%</td>
<td>18.2%</td>
</tr>
<tr>
<td>Multiple-Channel User</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>9.3%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Two Channel Buyer</td>
<td>8.9%</td>
<td>81.8%</td>
</tr>
<tr>
<td>Three Channel Buyer</td>
<td>0.4%</td>
<td>18.2%</td>
</tr>
<tr>
<td>No</td>
<td>91.7%</td>
<td>-</td>
</tr>
</tbody>
</table>

* “Mainly” means that at least 95% of purchases were made on that channel. Similarly, we classified as two channel users/three channel users those customers who made at least 95% of purchases using two channels/three channels.
### Table 5

Estimates of the Parameter Means– Multinomial Logit Channel Selection Switching Model

<table>
<thead>
<tr>
<th>Channel Choice Model Results&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Catalog vs Store</th>
<th>Internet vs Store</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter (sd)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Elasticity&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Intercept</strong></td>
<td><strong>Trial</strong></td>
<td>0.85 (0.47)</td>
</tr>
<tr>
<td></td>
<td><strong>Post-trial</strong></td>
<td>0.38 (0.92)</td>
</tr>
<tr>
<td><strong>Catalogs Sent</strong></td>
<td><strong>Trial</strong></td>
<td>-1.45 (0.36)</td>
</tr>
<tr>
<td></td>
<td><strong>Post-trial</strong></td>
<td>-0.12 (0.63)</td>
</tr>
<tr>
<td><strong>E-mails Sent</strong></td>
<td><strong>Trial</strong></td>
<td>-0.03 (0.61)</td>
</tr>
<tr>
<td></td>
<td><strong>Post-trial</strong></td>
<td>2.35 (0.34)</td>
</tr>
<tr>
<td><strong>State Dependence</strong></td>
<td><strong>Trial</strong></td>
<td><strong>4.09 (0.60)</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Post-trial</strong></td>
<td><strong>3.14 (0.62)</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup> Bold indicates that the 95% posterior interval for the parameter does not include zero.

<sup>b</sup> A positive coefficient means that a customer is more likely to choose channel j than the base channel. The base channel is the store.

<sup>c</sup> We computed elasticities at the mean value of the continuous variables and the modal of the categorical variables.
Table 6
Classifying Switchers as Having Preference-Based or Inertia-based Decision Processes

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Decision Strategy</th>
<th>Classification Rule&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Preference-Based</td>
<td>State Dependence &lt; Preference (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State Dependence &lt; Preference (I versus S)</td>
</tr>
<tr>
<td>2</td>
<td>Preference-Based</td>
<td>State Dependence &gt; Preference (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State Dependence &lt; Preference (I versus S)</td>
</tr>
<tr>
<td>3</td>
<td>Preference-Based</td>
<td>State Dependence &lt; Preference (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State Dependence &gt; Preference (I versus S)</td>
</tr>
<tr>
<td>4</td>
<td>Inertia-based</td>
<td>State Dependence &gt; Preference (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State Dependence &gt; Preference (I versus S)</td>
</tr>
</tbody>
</table>

<sup>a</sup> C stands for the catalog, I for the Internet and S for the store
Table 7

High Marketing versus Low Marketing Responsiveness

<table>
<thead>
<tr>
<th>Conditions&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Marketing Responsiveness&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Classification Rule&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HIGH</td>
<td>Mkt communications elasticity (C versus S) &gt; Median Mkt communications elasticity (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mkt communications elasticity (I versus S) &gt; Median Mkt communications elasticity (I versus S)</td>
</tr>
<tr>
<td>2</td>
<td>HIGH</td>
<td>Mkt communications elasticity (C versus S) &lt; Median Mkt communications elasticity (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mkt communications elasticity (I versus S) &gt; Median Mkt communications elasticity (I versus S)</td>
</tr>
<tr>
<td>3</td>
<td>HIGH</td>
<td>Mkt communications elasticity (C versus S) &gt; Median Mkt communications elasticity (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mkt communications elasticity (I versus S) &lt; Median Mkt communications elasticity (I versus S)</td>
</tr>
<tr>
<td>4</td>
<td>LOW</td>
<td>Mkt communications elasticity (C versus S) &lt; Median Mkt communications elasticity (C versus S)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mkt communications elasticity (I versus S) &lt; Median Mkt communications elasticity (I versus S)</td>
</tr>
</tbody>
</table>

<sup>a</sup> We have four conditions in total (three under High Marketing Responsiveness and one under Low Marketing Responsiveness) and consider two types of direct marketing communications (e-mails sent and catalogs sent). Therefore, we have a total of $4^2 = 16$ possible outcomes.

<sup>b</sup> We classify customers as low responsive to marketing only if all their marketing elasticities are less than the median values (or are not significant) across both the trial and post-trial models.

<sup>c</sup> C stands for the catalog, I for the Internet and S for the store.
Table 8

Migration Patterns – Channel Usage and Other Descriptive Information

<table>
<thead>
<tr>
<th>Code&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sample</th>
<th>Channel Usage—Percentage of times each channel is chosen in the trial and post-trial stages&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Average Order Size&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Average Amount Spent ($) &lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Percent</td>
<td>Trial (sd)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Post-Trial (sd)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>1→2</td>
<td>4.0</td>
<td>Mainly Catalog&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td>2.2 (0.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mainly Internet&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td>(0.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mainly Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>44.4% 11.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>33.3% 66.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog and Internet&lt;sup&gt;c&lt;/sup&gt;</td>
<td>- 22.2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internet and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog, Internet and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>22.2% -</td>
<td></td>
</tr>
<tr>
<td>3→1</td>
<td>42.2</td>
<td>Mainly Catalog&lt;sup&gt;c&lt;/sup&gt;</td>
<td>- 11.5%</td>
<td>2.1 (1.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mainly Internet&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td>(1.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mainly Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.1% 2.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>29.5% 34.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog and Internet&lt;sup&gt;c&lt;/sup&gt;</td>
<td>38.9% 46.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internet and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.2% -</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog, Internet and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>25.2% 5.26%</td>
<td></td>
</tr>
<tr>
<td>3→2</td>
<td>50.7</td>
<td>Mainly Catalog&lt;sup&gt;c&lt;/sup&gt;</td>
<td>- 0.8%</td>
<td>2.8 (2.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mainly Internet&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td>(3.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mainly Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>21.0% 14.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>44.7% 50.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog and Internet&lt;sup&gt;c&lt;/sup&gt;</td>
<td>24.6% 27.2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Internet and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>5.3% 3.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Catalog, Internet and Store&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.4% 4.4%</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Only migration patterns that represent more that 2% of the total are described.

<sup>b</sup> Simulation-based estimates across 100 iterations.

<sup>c</sup> Probability of choosing the same channel or a particular combination of channels over time greater than 95%.

<sup>d</sup> We computed the expected purchase occasion at which the switcher will switch as an average of simulation-based estimates across 100 iterations.
*By “theory” we mean the factors that have been identified as influencing customer’s channel choices. See Blattberg, Kim, and Neslin (2008) or Neslin et al. (2006) for discussion.
Figure 2

Channel Migrations Patterns from Trial to Post-Trial

TRIAL

Preference-Based

Inertial-Based

POST-TRIAL

High mkt?

yes

no

yes

no

yes

no

yes

no

Yes or no decision strategy change - Marketing change

Yes or no decision strategy change - Same Marketing

Stayers
Figure 3

Fraction of Customers Receiving E-mails and Catalogs over Time
Figure 4

Resulting Trial to Post-Trial Migrations Patterns

\(a\) Only paths greater than 2% threshold are represented
Figure 5  
*Switchers and Stayers* Channel Choice of the Internet in Response to Email Campaigns*

**A** *Switchers’* responsiveness to heavy-early versus heavy-late e-mail strategy**

---

**B** *Stayers’* responsiveness to heavy-early versus heavy-late e-mail strategy**

---

*Simulation-based estimates based on 100 iterations.*

**Two possible e-mail strategies were examined: heavy-early and heavy-late. Twenty purchase occasions were considered. In the heavy-early case, we schedule 70% of the average total observed e-mails sent per customer to the first five purchase occasions and 30% to the last fifteen occasions. In the latter case, we schedule 70% of these emails to the last five purchase occasions and 30% to the first fifteen occasions. "No marketing" means customers received no e-mails and no catalogs.*
APPENDIX: Convergence Diagnostics for the Parameter Means

History graphs parameter means for the intercepts in the trial stage

History graphs parameter means for the intercepts in the post-trial stage

History graphs parameter means for catalogs sent in the trial stage

History graphs parameter means for the catalogs sent in the post-trial stage
History graphs parameter means for e-mails sent in the trial stage

History graphs parameter means for e-mails sent in the post-trial stage

History graphs parameter means and Brooks-Gelman-Rubin plot for state dependence - trial stage

History graphs parameter means and Brooks-Gelman-Rubin plot for state dependence - post-trial stage

* For illustrative purposes we reported only the history graphs we obtained similar results also for the Brooks-Gelman-Rubin convergence statistic.