

# Decision Process Evolution in Customer Channel Choice

The growing number of sales channels through which customers can make purchases has made it imperative for managers to understand how customers decide which channels to use. However, this presents a significant challenge because there is reason to believe the channel decision process evolves over the lifetime of the customer. The authors document the existence and nature of this phenomenon by analyzing the evolution of a customer's channel choice decision process from a trial stage to a posttrial stage. First, they analyze data for a book retailer and replicate their analysis using data from a durables and apparel retailer. Their results suggest that (1) customers' decision processes do evolve, (2) a minority but sizeable segment changes decision processes within the observation period, and (3) customers who change do so from a decision process in which they are highly responsive to marketing to one in which they are less responsive. The authors illustrate and discuss the implications for both managers and researchers.

**Keywords:** multichannel marketing, channel migration, customer relationship management, customer channel choice

The ever-expanding multiplicity of channels through which customers can purchase makes it imperative for managers to understand how customers decide which channels to use. One potential benefit of this understanding is that managers can then fine-tune channel strategies to enhance customer value through effective multichannel customer management (Neslin et al. 2006). While researchers have made progress in understanding the customer channel decision process (Blattberg, Kim, and Neslin 2008, chap. 25), the managerial press reports that retailers' knowledge of this decision process is still limited and that surprisingly few companies are effective in managing multichannel strategies (Myers, Van Metre, and Pickersgill 2004; Noble, Shenkan, and Shi 2009; Weinberg, Parise, and Guinan 2007).

We propose that one factor contributing to the difficulty in understanding the customer channel decision process is that the process is a moving target—in other words, it evolves over the lifetime of the customer. For example, this would mean that managers need to treat new-to-the-firm

customers differently than mature customers in terms of multichannel customer policy. A newly acquired customer may be in a “deliberative” mind-set (Gollwitzer and Bayer 1999) and thus be open to suggestions about which channel to use (i.e., responsive to marketing efforts); in contrast, as the customer learns more about the firm's channels, he or she becomes less receptive to these efforts. This would mean that the window during which to “right-channel” the customer (Neslin and Shankar 2009) may be directly after the customer is acquired.

Therefore, it is critical that managers understand if, and if so how, the customer channel decision process evolves. The purpose of this study is to investigate the existence and nature of this evolution and explore its managerial implications. Our focus is not about *which* channels consumers choose to make their purchases in, but *how* they decide which channels to choose for purchase and how that decision process changes over time. We examine four major questions:

1. Does the customer channel choice decision process change over time? That is, is decision process evolution a real phenomenon?
2. If so, how many customers “learn” a new decision process quickly (learners) versus how many customers evolve more gradually (stayers)?
3. How does the learners' initial, or trial, stage decision process compare with their later or posttrial stage decision process?
4. What are the profit implications for more effective multichannel customer management?

Our findings across two data sets confirm that the customer decision process changes over time. In a data set in which a book club retailer acquires customers and we observe these customers over time, a significant learner segment (22% of customers) becomes less inertial with respect

---

Sara Valentini is Assistant Professor of Marketing (e-mail: s.valentini@unibo.it), and Elisa Montaguti is Associate Professor of Marketing (e-mail: elisa.montaguti@unibo.it), Management Department, University of Bologna. Scott A. Neslin is Albert Wesley Frey Professor of Marketing, Tuck School of Business, Dartmouth College (e-mail: scott.a.neslin@dartmouth.edu). The authors thank Leonard Lee and Carl Mela for helpful discussions on behavioral theory and statistical analysis, respectively; Bruce Hardie for comments and suggestions; and Paul Wolfson for computer programming support. They thank two anonymous companies that provided the data used in this research. Finally, they express their appreciation to the three anonymous *JM* reviewers for their diligence and constructive suggestions.

---

to channel choices and more driven by channel preferences. At first, their channel decisions are shaped significantly by the firm's marketing efforts, but later they become less responsive to marketing. These changes are set in motion soon after a company "acquires" a customer. Therefore, companies should try to "right-channel" customers soon after they are acquired. Customers who are more likely to learn can be characterized according to directly measurable indicators such as how these customers are acquired, their proclivity to provide their e-mail address, and their age. We illustrate that by using these characteristics to identify the learner segment and by using e-mail to transition them early into the cheaper Internet channel, the company can increase its profits by approximately 24%. In a second data set, provided by a durables/apparel retailer that has just introduced a new purchase channel, we confirm that customers become less responsive to marketing over time, although do not find that the process becomes more preference driven over time.

We proceed as follows: We develop a framework for the channel decision process and a taxonomy of potential evolution patterns. Then, we discuss the theory of why decision processes should evolve and advance hypotheses. Next, we describe our methodology for measuring decision process evolution. We then show our data, results, and profitability illustration. We conclude with a discussion of implications for research and practice.

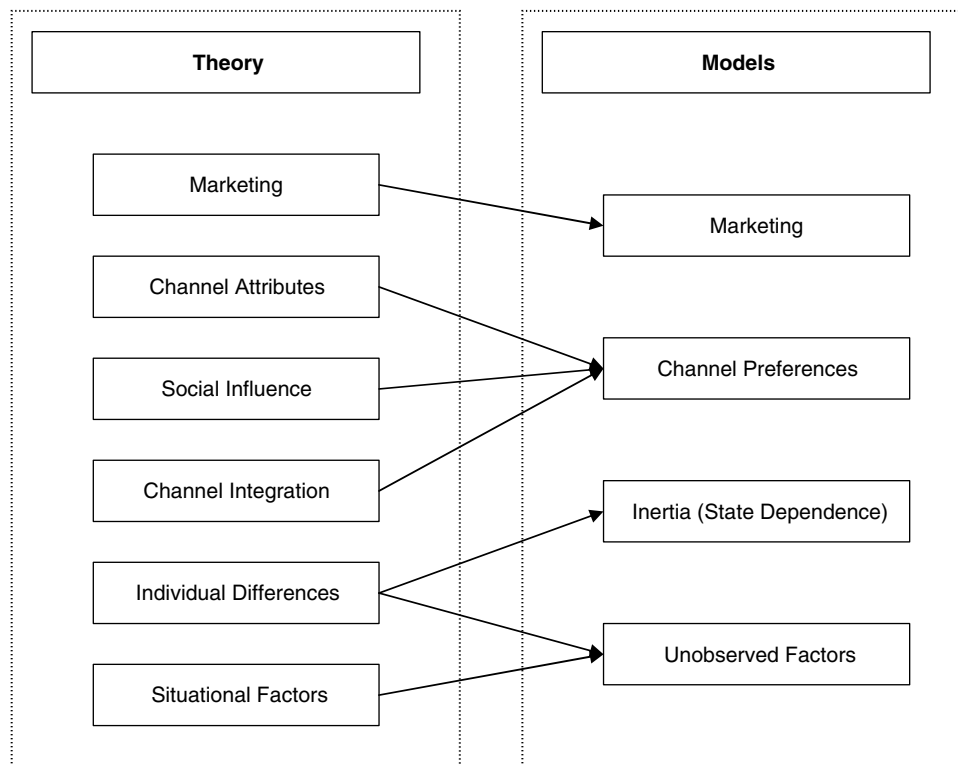
## A Taxonomy for Categorizing Decision Process Evolution

Blattberg, Kim, and Neslin (2008) and Neslin et al. (2006) summarize six factors that determine customer channel choice: (1) marketing, (2) channel attributes, (3) social influence, (4) channel integration, (5) individual differences, and (6) situational factors (see Figure 1). Researchers have organized these factors into marketing, channel preferences, inertia or state dependence, and unobserved factors, as shown in the right side of Figure 1. This suggests that the channel choice decision can be represented as follows:

$$(1) \quad \text{Channel choice} = f(\text{marketing, channel preference, state dependence}).$$

The drivers of channel choice in Equation 1 and their relative impact characterize the channel choice decision process. Several researchers have estimated models according to Equation 1 (Ansari, Mela, and Neslin 2008; Knox 2006; Thomas and Sullivan 2005; Venkatesan, Kumar, and Ravishanker 2007; Verhoef, Neslin, and Vroomen 2007). Knox shows that channel preference can change over time. A common finding is that customers differ in the relative impact the three drivers have on their channel choice process. We account for this heterogeneity and also allow the relative impact of the three factors to evolve over time.

**FIGURE 1**  
**A Framework for the Multichannel Choice Process: Theory Translated into Models**



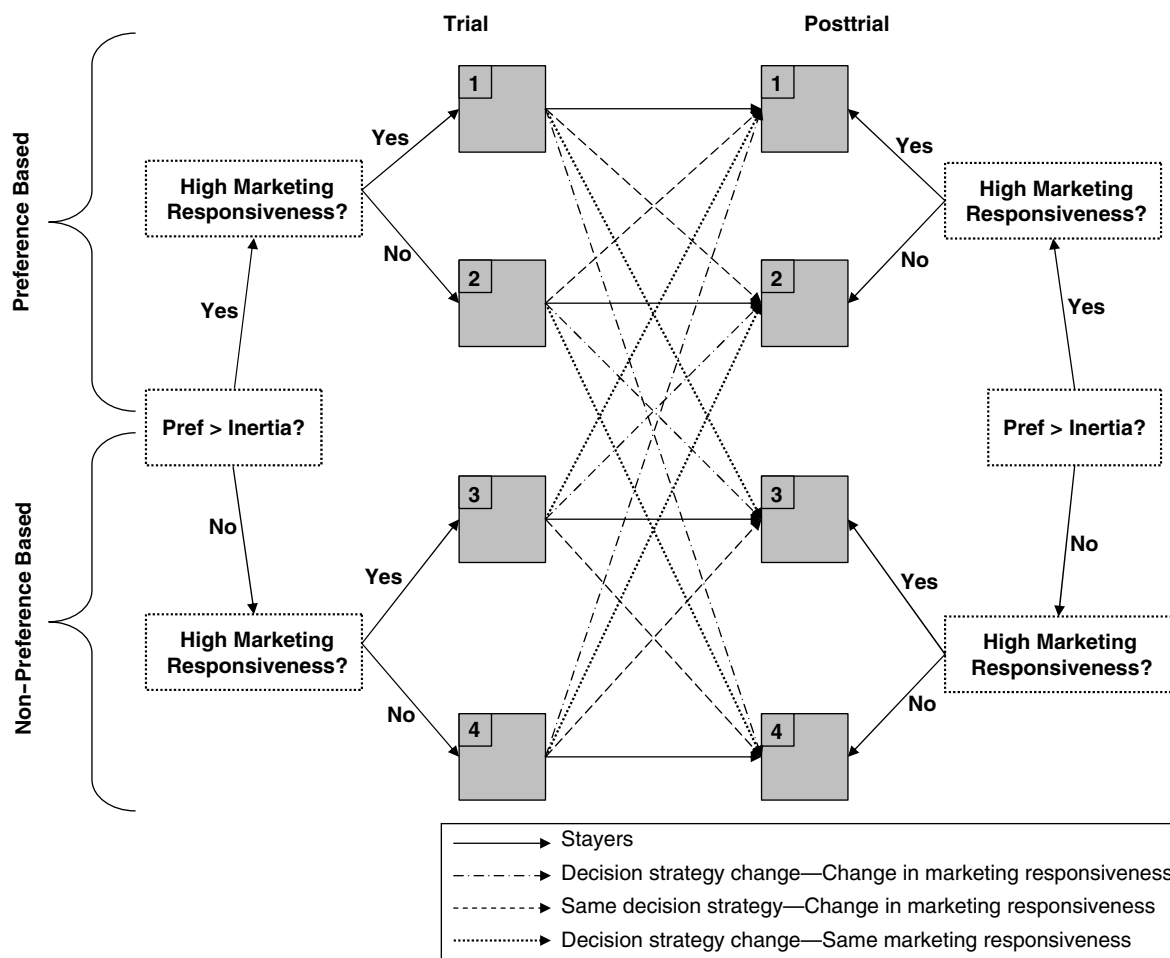
Notes: By "theory," we mean the factors that have been identified as influencing customer's channel choices (for a discussion, see Blattberg, Kim, and Neslin 2008; Neslin et al. 2006).

Figure 2 shows how the relative impact of preferences, marketing, and inertia define a taxonomy of potential evolution patterns. Starting from the left, we ask whether the decision is preference based or non-preference based. Next, we ask whether the customer is marketing responsive or not marketing responsive. This produces four potential decision processes. The arrows show how the customer's process might evolve from a trial to a posttrial stage. A customer's trial stage process might be non-preference based but highly responsive to marketing (Category 3), but it might evolve to a posttrial classification of preference based and not marketing responsive (Category 2). There are 12 possible evolution patterns and four cases in which the process does not change.

## Hypotheses

This first question is, Why should the decision process evolve? Theory often starts with observation, and this is the case with decision process evolution. Aaker (1971) finds that preferences for a new product develop over a trial stage, in which the consumer learns about the product, to a post-trial stage, in which the consumer has learned his or her final preference for the product and purchases accordingly. This can be viewed as isomorphic to Gollwitzer and Bayer's (1999) theory that consumer decision making moves from a deliberative to an implemental stage. In the deliberative stage, the consumer is trying to determine what to do, which is then executed during the implemental stage. Aaker's and

**FIGURE 2**  
A Taxonomy of Channel Migration Patterns from Trial to Posttrial



Notes: 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 are less influenced by marketing activities. For example, they always buy on the Internet, and marketing has little influence on which channel they choose. Customers in Category 3 tend to use the channel they used the previous 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 induces them to use a different channel. 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. The figure also shows the possibility that the process does not change; that is, the customer is a stayer. This is signified by the solid lines in the figure.

Gollwitzer and Bayer's work suggests two stages of decision making, and thus we adopt a two-stage framework (Figure 2). However, more stages could be hypothesized (Fader, Hardie, and Huang 2004; Heilman, Douglas, and Wright 2000), and this could be an avenue for further research.

The theoretical concept at work in decision process evolution is fundamentally one of learning. What is being learned is the decision process (i.e., how the customer will go about choosing which channel to use to make a purchase). Therefore, learning means that the customer determines his or her channel preferences, how he or she will respond to the firm's marketing communications, and to what extent he or she will simplify things by using the same channel used previously. In terms of Gollwitzer and Bayer's (1999) theory, the customer first adopts a deliberative mind-set, determining what his or her decision process will be. After this process is learned, the customer adopts an implemental mind-set, executing the chosen decision process.

This clarifies *what* is being learned—the decision process—which in turn prompts the question of *why* it is being learned. There are at least three reasons:

1. *Motivation and ability*: Motivation and ability are prerequisites to learning (Bettman and Park 1980; Hoch and Deighton 1989; MacInnis, Moorman, and Jaworski 1991). These factors are salient when the cost of an incorrect decision is high. For example, the customer may buy a product for a specific event using a catalog, but if the product is delivered late, the customer pays a cost by not having it when it was needed. These factors are also relevant when customers have the ability to evaluate their decisions. The channel experience is easy to evaluate in terms of service convenience, effort, and so on (Verhoef, Neslin, and Vroomen 2007).
2. *Lack of task familiarity*: Learning is more likely to take place when customers are not familiar with the task at hand (Alba and Hutchinson 1987; Hoch and Deighton 1989; Johnson and Russo 1984). For example, newly acquired customers may have experience with buying on the Internet or with the policies of certain stores but not with the channels offered by the particular company for which they have just become a customer. Thus, when they are not familiar with the channel environment, there are important things for them to learn.
3. *Unsatisfying experiences*: An unsatisfying experience (e.g., a channel was not as convenient as expected) may cause the customer to reassess and try another channel. Previous research has shown that unsatisfying experiences can have a profound impact on future behavior (Mattila 2003; Weiner 1986).

In summary, a customer's decision processes should evolve because of learning, to the extent that motivation and ability, lack of task familiarity, and unsatisfying experiences are present.<sup>1</sup> These factors can be set in motion in several contexts (e.g., directly after the customer has been acquired, when there is an abrupt change in the environment; Moe and Yang 2009). Therefore, we advance the following hypothesis:

---

<sup>1</sup>Note that although we emphasize the role of learning, there are other potential reasons the decision process might change. In the channel choice context, these would include obtaining better Internet connectivity.

H<sub>1</sub>: The channel choice process for the average customer evolves over time.

We specify "average customer" because the factors driving learning—motivation and ability, lack of task familiarity, and unsatisfying experience—vary across customers. Some customers may be uninvolved shoppers and thus not motivated to evaluate their channel experience. Some (e.g., those who shop for others' in addition to their own needs) may not fully be able to evaluate the outcome of their channel choice. Others may be so experienced with shopping that they are highly familiar with the channel selection process. Some will have unsatisfying experiences; others will not. As a result, we hypothesize the following:

H<sub>2</sub>: The likelihood that the channel decision process evolves varies across customers.

We now turn to hypothesizing the patterns of evolution most likely to surface when we classify customers into the taxonomy depicted in Figure 2. The literature suggests the trial stage will be relatively less driven by preferences than the posttrial stage (Heilman, Douglas, and Wright 2000; Meyer and Sathi 1985). Gollwitzer and Bayer (1999) suggest that when customers are in a deliberative mind-set, they are uncertain about their goals. In the channel choice context this means that they are unsure of what channel attributes are relevant and how channels rate on those attributes. This suggests that consumers are less (more) certain of their channel preferences when they are in the trial (posttrial) stage. This leads to the following hypothesis:

H<sub>3</sub>: The trial stage of the customer decision process is less preference dominated than the posttrial stage.

The question of which channel to choose is inherently more challenging for a newly acquired customer or a customer facing a new channel environment. Therefore, the trial stage of the customer decision process is characterized by complexity as well as goal uncertainty. Consumers are known to resort to simplifying heuristics in such situations (Payne, Bettman, and Johnson 1992; Tversky and Kahneman 1974). One possible heuristic, essentially the availability heuristic (Tversky and Kahneman 1974), would be to bias current channel choices toward the channel last chosen. This would apply to the trial stage. However, as consumers make more choices and become more familiar with the channel choice task, that task becomes less complex and consumers develop clearer goals. This suggests that consumers are more (less) likely to rely on heuristics and inertia in the trial (posttrial) stage. Therefore, we propose the following:

H<sub>4</sub>: The trial stage of the customer decision process is more inertia dominated than the posttrial stage.

Gollwitzer and Bayer (1999) theorize that consumers in the deliberative stage have an open mind-set and are responsive to information. When they move to the implemental stage, they have less need for information, so communications are less influential. Previous work has explored temporal differences in marketing effects. Narayanan, Manchanda, and Chintagunta (2005) find that marketing communications reduce uncertainty about pharmaceutical product quality in the early launch phases and reinforce preferences in

subsequent stages. Overall, information has a lower impact over time (see Figure 1 in Narayanan, Manchanda, and Chintagunta 2005). Similarly, Lee and Ariely (2006) show that promotion effectiveness varies depending on the consumer's mind-set (i.e., deliberative vs. implemental) and, in particular, that it is greater when consumers' goals are less well defined. These considerations suggest the following hypothesis:

H<sub>5</sub>: The trial stage should be more marketing responsive than the posttrial stage.

## Methodology

Following our goals, we sought to develop a methodology for determining whether the customer channel choice decision process evolves over time, how this differs across customers, and how the trial stage process compares with the posttrial stage process. To accomplish this, we build on previous work (Ansari, Mela, and Neslin 2008; Thomas and Sullivan 2005) and develop an analysis consisting of two components: (1) a model that captures the evolution process and (2) two logit channel choice models, one for the trial stage and the other for the posttrial stage. All parameters in these models vary across customers, so we can examine how the evolution process differs from customer to customer and distinguish fast evolvers (learners) and slower evolvers (stayers).

The evolution component quantifies how many purchase occasions the customer takes to transition from the trial to the posttrial stage. We use the purchase occasion as our unit of analysis because our theory posits that the decision process changes as a result of learning from the customer's channel experiences. Specifically, we analyze the number of purchase occasions it takes for the customer to evolve from the trial to the posttrial stage using a simple geometric distribution. Specifically, let  $q_h$  be the conditional probability on any given purchase occasion customer  $h$  moves from the trial to the posttrial stage, given the customer has not moved yet, and let  $p_{ht}$  be the probability customer  $h$  has moved to the posttrial stage by the  $t$ th purchase occasion. These two quantities are related as follows:

$$(2) \quad p_{ht} = 1 - (1 - q_h)^{t-1}.$$

This is because  $(1 - q_h)^{t-1}$  is the probability the customer has not moved by the  $t$ -th purchase occasion, so the converse,  $1 - (1 - q_h)^{t-1}$ , is the probability the customer has moved by the  $t$ -th purchase occasion. Note that  $p_{ht}$  is an increasing function of  $q_h$  (the more likely the customer is to change on any given purchase occasion, the more likely it is that the customer will have changed by purchase occasion  $t$ ) and an increasing function of  $t$  (the more purchase occasions, the more likely it is that the customer will have evolved).

The geometric distribution with a single parameter ( $q_h$ ) is parsimonious, though it implies that the most likely time for the customer to evolve is immediately after acquisition or after the decision environment changes. To investigate this assumption, we tested two alternative distributions (the discrete Weibull and a conditional logit). The results show that the geometric specification outperformed the others (see Web Appendix W1 at <http://www.marketing>

[.com/jmnov11](http://www.marketing.com/jmnov11)). We also note that the estimated values of  $q_h$  suggest that the mean time customers take to evolve to the posttrial stage is 23 purchase occasions, with a standard deviation of 5.4. This means there is much variability in evolution rates, and by no means do most customers evolve right away.

The second component of our analysis consists of two logit channel choice models, one for the trial stage and one for the posttrial stage. The term  $U_{hjt}$  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 indicates posttrial. Therefore,  $U_{hjt}^0$  is the utility for customer  $h$  in choosing channel  $j$  at purchase occasion  $t$ , given the customer is in the trial stage, and  $U_{hjt}^1$  is the utility of choosing channel  $j$  at purchase occasion  $t$ , given the customer is in the posttrial stage. These utilities are a function of the factors shown in Equation 1 and Figure 1:

$$(3) \quad U_{hjt}^0 = \alpha_{hj}^0 + \beta_{1hj}^0 CS_{ht} + \beta_{2hj}^0 ES_{ht} + \beta_{3h}^0 LC_{hjt} + \varepsilon_{hjt}^0, \quad \text{and}$$

$$(4) \quad U_{hjt}^1 = \alpha_{hj}^1 + \beta_{1hj}^1 CS_{ht} + \beta_{2hj}^1 ES_{ht} + \beta_{3h}^1 LC_{hjt} + \varepsilon_{hjt}^1,$$

where

- $\alpha_{hj}^k$  = customer  $h$ 's preference for channel  $j$  during the trial stage ( $k = 0$ ) or the posttrial stage ( $k = 1$ );
- $CS_{ht}$  = number of catalogs sent to customer  $h$  during the quarter in which purchase occasion  $t$  occurs;
- $ES_{ht}$  = number of e-mails sent to customer  $h$  during the quarter in which purchase occasion  $t$  occurs;
- $LC_{hjt}$  = 1 if customer  $h$  purchased from channel  $j$  on the purchase occasion before purchase occasion  $t$  and 0 if otherwise. This captures state dependence;<sup>2</sup>
- $\beta_{mhj}^k$  = impact of catalogs ( $m = 1$ ) or e-mails ( $m = 2$ ) on customer  $h$ 's utility for channel  $j$  during the trial stage ( $k = 0$ ) or the posttrial stage ( $k = 1$ );
- $\beta_{3h}^k$  = impact of the previous purchase being in a particular channel on the utility for that channel, for customer  $h$  at purchase occasion  $t$ , during the trial stage ( $k = 0$ ) or the posttrial stage ( $k = 1$ ); and
- $\varepsilon_{hjt}^k$  = error term for customer  $h$  for channel  $j$  during the trial stage ( $k = 0$ ) or the posttrial stage ( $k = 1$ ) for purchase occasion  $t$ . For example, this includes situational factors not directly observable by the researcher.

The term  $\alpha_{hj}^k$  captures channel preferences,  $\beta_{mhj}^k$  captures the importance of marketing, and  $\beta_{3h}^k$  captures the importance of inertia or state dependence. The relative sizes of these customer-specific parameters enable us to classify customers into the categories defined in Figure 2, both before and after they change decision processes.

Using Equations 3 and 4 and assuming an extreme value distribution for the error terms, we have two logit choice

<sup>2</sup>State dependence refers to the influence of previous choices on current choices. We used a simple zero-one formulation of the state dependence variable (LC) because previous research is dominated by the operationalization of state dependence to equal 1 if the alternative was chosen at the previous purchase occasion (e.g., Ansari, Mela, and Neslin 2008; Thomas and Sullivan 2005).

models. Equation 2 indicates the likelihood the customer is using the trial or posttrial model at purchase occasion  $t$ . Combining these components yields the following probability customer  $h$  chooses channel  $j$  at purchase occasion  $t$ :<sup>3</sup>

$$(5) P_{hjt} = (1 - p_{ht}) \left( \frac{\exp(U_{hjt}^0)}{\sum_{j=1}^J \exp(U_{hjt}^0)} \right) + p_{ht} \left( \frac{\exp(U_{hjt}^1)}{\sum_{j=1}^J \exp(U_{hjt}^1)} \right).$$

With probability  $(1 - p_{ht})$ , the customer is still in the trial stage and thus is using the first logit model; with probability  $p_{ht}$ , the customer has evolved and is using the second logit model. On the customer's first purchase occasion,  $t = 1$ , so  $p_{ht} = 0$  (Equation 2), and the customer's decision process is driven entirely by the trial model. Over time,  $p_{ht}$  increases to a degree determined by the customer's parameter  $q_h$ , and the process becomes increasingly driven by the posttrial model. For any purchase occasion, the customer's probability of choosing channel  $j$  is a weighted average of the trial and posttrial models, the weights being the likelihood that the customer is in the trial stage or the posttrial stage, respectively. Because the weight placed on the posttrial model increases as the customer accumulates more purchases (Equation 2), the posttrial model becomes more important as purchases accumulate.

## Data and Estimation

Our main data set is from a major European book retailer operating in one country. The retailer uses three sales channels—physical retail stores, catalogs, and the Internet—and uses a subscription business model. Thus, to be eligible to purchase, each customer must become a member. A code is associated with each customer, tracking each time she or he purchases and from which channel. Therefore, we have information on which channel each customer selected on each purchase occasion, the date of each purchase, how much was spent, the number and types of communications (e-mails and catalogs) each customer received and when they were received, and age and gender demographics.

Each channel shares the same assortment and price. Communications were not targeted according to channel usage. This means that each channel was treated the same in terms of marketing activities. This gives us the opportunity to measure intrinsic customer preferences for each channel.

The period of observation is January 2002–June 2006. The retailer monitored customers and made customer-marketing decisions on a quarterly basis. Therefore, we

<sup>3</sup>We also estimated Equation 5 with interactions effects and/or diminishing returns to marketing investments (e.g., Thomas and Sullivan 2005; Venkatesan, Kumar, and Ravishanker 2007). In addition, we tested a model with a catalog stock variable that allows for cumulative and carryover effects as the customer gets more catalogs. However, the fit of all these models was lower, and the interpretation of the parameters estimates was essentially identical.

assemble each customer's purchase occasions, channel choice, and the catalogs and e-mails received on a quarterly basis. On the rare occasions when there was more than one purchase occasion in a quarter (this happened in only 4.2% of the 14,985 customer-quarter observations), we used the single purchase occasion. We found no difference in the results when we estimated the model using all purchase occasions.

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 fourth quarter of 2001. We focus on a static sample of active customers (i.e., customers having an active relationship with the company between January 2002 and June 2006) with all three channels available throughout the relationship. The sample size is 1018 households. Table 1 presents descriptive information about the sample. We use Bayesian estimation (see Web Appendix W2 at <http://www.marketingpower.com/jmnov11>) to estimate the model.

Catalogs and e-mails sent ( $CS_{ht}$  and  $ES_{ht}$ ) can depend on the customer's past choices and thus might be endogenous. We used an approach similar to Gönül, Kim, and Shi (2000) to minimize endogeneity bias (see Web Appendix W3 at <http://www.marketingpower.com/jmnov11>).

**TABLE 1**  
**Descriptive Statistics of the Selected Cohort of Customers (n = 1018)**

Variable	M	SD	Min	Max
Average number of purchase occasions (per year)	3.1	1.1	.0	4.0
Average number of purchase occasions over relationship	14.7	2.9	7.0	18.0
Average returns (in US\$ per quarter)	.4	3.5	.0	104.4
Average number of catalogs received (per quarter)	2.0	.7	.0	4.0
Average number of e-mails received (per quarter)	.7	2.0	.0	12.0
Age (years)	43.8	14.6	21.0	88.0
Gender (male)	35.2%			

Channel Usage <sup>a</sup>	% of Customers	Average Number of Purchase Occasions per Customer
Mainly catalog	27.6%	14.6
Mainly Internet	.6%	14.8
Mainly store	42.4%	15.3
Catalog and store	15.4%	13.4
Catalog and Internet	8.4%	14.7
Internet and store	1.2%	15.3
Catalog, Internet, and store	4.3%	14.6

<sup>a</sup>"Mainly" means that at least 95% of purchases were made on that channel. Similarly, we classified as two-channel/three-channel users those customers who made at least 95% of purchases using two channels/three channels.

## Results

### *The Existence of Choice Evolution and Its Variation Across Customers*

We test for the existence of decision process evolution ( $H_1$ ) by comparing our proposed model (Equations 2–5) with a logit model that does not distinguish between trial and posttrial stages (Model 1).<sup>4</sup> We test whether the likelihood of evolution varies across customers ( $H_2$ ) by comparing the proposed model with a logit model that assumes each customer starts off using one logit model and moves to another logit model after the same a priori defined number of purchase occasions (nine; Model 2).<sup>5</sup> We compare these models using both the deviance information criterion (DIC) statistic (Spiegelhalter et al. 2002) and the hit rate (see Web Appendix W4 at <http://www.marketingpower.com/jmnov11>). The proposed model (DIC = 6525.2, hit rate = 72.9%), which assumes a distinct trial and posttrial logit model and a heterogeneous trial period length across customers, is superior to both Model 1 (DIC = 6881.0, hit rate = 57.9%) and Model 2 (DIC = 6648.2, hit rate = 56.9%). These results support  $H_1$  and  $H_2$ , indicating the existence of channel choice evolution and suggesting that the likelihood of evolution varies across customers.

### *Prevalence of Decision Process Evolution and Channel Usage of Learners Versus Stayers*

The preceding results suggest that customers change decision processes. Next, we want to know how prevalent this behavior is. To investigate this, we compute the probability that customer  $h$  evolves to the posttrial stage by his or her last observed purchase occasion (see Web Appendix W5 at <http://www.marketingpower.com/jmnov11>). The expected value of this probability is .22, which means that we expect 22% of customers to have changed decision processes by the end of the observation period. We ordered customers according to their individual probability and labeled the 22% highest as the learners. The logic is that the customers with the highest probability of changing are the ones we would expect to change by the end of the observation period. The remaining customers were classified as stayers (i.e., not expected to change by the end of the observation period). The average probability of changing for stayers is .14 and for learners is .49. Using Equation 2, we calculated that, on average, it takes the learner 9 purchase occasions to evolve (see Web Appendix W5 at <http://www.marketingpower.com/jmnov11>). By the same token, it takes

<sup>4</sup>Model 1 essentially assumes that  $\alpha_{hj}^0 = \alpha_{hj}^1$  and  $\beta_{mhj}^0 = \beta_{mhj}^1$  for  $m = 1, 2, 3$  in Equations 3–4.

<sup>5</sup>Nine is the median number of purchase occasions. We tested other numbers and also found the proposed model was superior (6715.9 for 3 purchase occasions, 6725.5 for 5, 6804.3 for 12, and 6970.6 for 15 versus DIC = 6525.8 for the proposed model). The model with 1 purchase occasion yielded DIC = 6631.1 but did not converge. We also tested a model for which  $q_h$  was constant across customers (DIC = 10,204.0), but it also did not converge.

an average of 32 purchase occasions for a stayer to evolve, indicating a clear separation between the two groups.

Table 2 shows learners' versus stayers' usage of various channels and reveals an intriguing finding: Stayers are mainly single-channel users, while learners are predominantly multichannel. We conducted a sensitivity analysis regarding this finding by relaxing the 95% requirement used to identify single channel customers (see note a in Table 2). We find that this result is robust to the rule for defining a multichannel customer. This is an important exploratory finding and a characterization of the multichannel customer previously not demonstrated.

### *Describing and Contrasting the Trial and Posttrial Decision Processes*

Table 3 presents the estimates of the mean parameters for the proposed model. The intercept estimates suggest that customers prefer the store to the Internet in both trial and posttrial stages. This may seem counterintuitive given the popularity of the Internet in book retailing. However, the customers in our sample essentially had joined a loyalty club for a book retailer and therefore presumably were avid readers. It is sensible that such customers would enjoy browsing through books and purchasing in the store. Also, the stores were a relatively new channel for the retailer, which opened its first store in 2000. Thus, in contrast to the prevailing tradition in the book industry of bookstores being the "incumbent" channel, in our case, the store was the "novel" channel. While the store is, on average, preferred over the Internet, the intercept for the catalog over the store is not significant for either the trial or posttrial periods, showing that these two channels were equally preferred on average.

We now contrast the trial and posttrial decision processes. We tested whether differences between the trial and posttrial coefficients were statistically significant (see

**TABLE 2**  
**Stayers Versus Learners: Channel Choice Behavior<sup>a</sup>**

	Stayers (793 Customers)	Learners (225 Customers)
<b>Channel Usage</b>		
Mainly catalog	35.4%	—
Mainly Internet	.8%	—
Mainly store	54.5%	—
Catalog and store	6.2%	48.0%
Catalog and Internet	2.6%	28.9%
Internet and store	.1%	4.9%
Catalog, Internet, and store	.4%	18.2%
<b>Multiple-Channel Shopper</b>		
Yes	9.3%	100.0%
Two-channel buyer	8.9%	81.8%
Three-channel buyer	.4%	18.2%
No	90.7%	—

<sup>a</sup>"Mainly" means that at least 95% of purchases were made on that channel. Similarly, we classified as two-channel/three-channel users those customers who made at least 95% of purchases using two channels/three channels.

**TABLE 3**  
**Estimates of Parameter Means: Two-Stage Channel Choice Model**

	Catalog Versus Store		Internet Versus Store	
	Parameter <sup>a</sup>	Elasticity <sup>b</sup>	Parameter <sup>a</sup>	Elasticity <sup>b</sup>
<b>Intercept</b>				
Trial	.85 (.47)	—	<b>-2.48</b> (.39)	—
Posttrial	.38 (.92)	—	<b>-3.78</b> (.46)	—
<b>Catalogs Sent</b>				
Trial	<b>-1.45</b> (.36)	<b>-7.24</b>	<b>-9.10</b> (.97)	<b>-20.56</b>
Posttrial	-.12 (.63)	-.33	.63 (.59)	1.83
<b>E-mails Sent</b>				
Trial	-.03 (.61)	-.46	<b>3.28</b> (.75)	<b>4.90</b>
Posttrial	<b>2.35</b> (.34)	<b>.93</b>	.41 (.70)	-.49
<b>State Dependence</b>				
Trial		<b>4.09</b> (.60)		
Posttrial		<b>3.14</b> (.62)		

<sup>a</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>b</sup>We computed elasticities at the mean value of the continuous variables and the mode of the categorical variables.

Notes: Bold indicates that the 95% posterior distribution for the parameter does not include zero. The standard deviation of the posterior distribution is in parentheses. For identification purposes, we set one channel (the store) as the base.

Web Appendix W6 at <http://www.marketingpower.com/jmnov11>). We found that the state dependence effect is significantly reduced in the posttrial period, suggesting that, on average, people become less inertial. By a similar token, we found that preferences for the store versus the Internet significantly increase. This suggests that in the post-trial stage, customers base their decision processes more on channel preferences than habitual behavior, as we predict in  $H_3$  and  $H_4$ .

We have four marketing coefficients for each stage (catalog's impact on catalog vs. store and Internet vs. store, and e-mail's impact on catalog vs. store vs. Internet vs. store). In three of four cases, the marketing coefficient becomes significantly smaller in the posttrial model, meaning that customers evolve toward a decision process in which they have less need for marketing information, in support of  $H_5$ .

The preceding conclusions are based on comparing the magnitudes of coefficients. We also examined pseudo *t*-statistics, calculated as the mean value of the coefficient divided by its standard deviation (LeSage and Fischer 2010, p. 429). These results are consistent with the magnitude results noted previously (the Internet vs. store preference pseudo *t*-statistic increases from  $2.48/.39 = 6.36$  in the trial period to  $3.78/.46 = 8.22$  in the posttrial period, the state dependence statistic decreases from 6.82 to 5.06, and the pseudo *t*-statistics for marketing decrease in the three cases in which the magnitudes decrease and increase in the one case in which the magnitude increases).

In summary, these results suggest that the decision process evolves in two main ways: (1) Channel preferences become more important and state dependence less important over time ( $H_3$  and  $H_4$ ), and (2) in general, marketing communication becomes less effective ( $H_5$ ). It is noteworthy that our findings indicate consumers do not always link catalogs and e-mails to their "natural" purchase channels, the catalog and Internet, respectively. For example, the negative sign for the catalog's impact on catalog versus store suggests that catalogs "promote" the store. The catalogs used by this company did not include store promotions but did list

the locations of stores. This could account for the tendency of catalogs to promote store choice over catalog in the trial period. Another possible explanation is the "research shopper" phenomenon (Verhoef, Neslin, and Vroomen 2007). Research shopping means the customer uses one channel for search (gathering information) and a different channel for purchase. For example, customers might find catalogs to be a convenient way to investigate what products are available, but the store is a superior purchase channel for them because they can sample the product by thumbing through the books. We also found that e-mails promote catalogs versus store in the posttrial period. It may be that readers of e-mails are goal-directed and thus gather all the information they need from the e-mail. They then find it more convenient to purchase the product from the catalog, which just requires a phone call, than from the store. Blattberg, Kim, and Neslin (2008, p. 647) articulate a broader view of this by raising the notion of channel/marketing congruency. They state that this is an underresearched area, and although it is not the main purpose of this article, our results add to the substantive knowledge base on this issue.

### **Characterizing Evolution Patterns**

We now focus on learners and classify them in terms of the taxonomy depicted in Figure 2; in other words, we quantify the extent to which different decision processes are employed in the trial versus posttrial stages and examine the patterns by which learners evolve from one process to another. The customer-level parameter estimates enable us to accomplish this classification.

By comparing the magnitudes of the channel preference parameters with the state dependence parameters, we classify customers' decision processes as either preference based or non-preference based. For example, if customer *h*'s state dependence parameter is higher than his or her preference for the catalog over the store and the Internet over the store (in absolute values), we classify this customer as non-preference based (for details, see Table 4, Panel A).



**TABLE 4**  
**Classifying Learners**

<b>A: Preference Based or Non-Preference Based</b>		
<b>Categories</b>	<b>Decision Strategy</b>	<b>Classification Rule<sup>a</sup></b>
1	Preference based	State dependence < Preference (C vs. S) State dependence < Preference (I vs. S)
2	Preference based	State dependence > Preference (C vs. S) State dependence < Preference (I vs. S)
3	Preference based	State dependence < Preference (C vs. S) State dependence > Preference (I vs. S)
4	Non-preference based	State dependence > Preference (C vs. S) State dependence > Preference (I vs. S)
<b>B: High Marketing Versus Low Marketing Responsiveness</b>		
<b>Categories<sup>b</sup></b>	<b>Marketing Responsiveness<sup>c</sup></b>	<b>Classification Rule<sup>a</sup></b>
1	High	Marketing elasticity (C vs. S) > Median marketing elasticity (C vs. S) Marketing elasticity (I vs. S) > Median marketing elasticity (I vs. S)
2	High	Marketing elasticity (C vs. S) < Median marketing elasticity (C vs. S) Marketing elasticity (I vs. S) > Median marketing elasticity (I vs. S)
3	High	Marketing elasticity (C vs. S) > Median marketing elasticity (C vs. S) Marketing elasticity (I vs. S) < Median marketing elasticity (I vs. S)
4	Low	Marketing elasticity (C vs. S) < Median marketing elasticity (C vs. S) Marketing elasticity (I vs. S) < Median marketing elasticity (I vs. S)

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

<sup>b</sup>We have four categories 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>c</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 posttrial models.

After classifying customers as preference or non-preference based, we further distinguish them according to their marketing responsiveness. We compare individual marketing elasticities with the median value across customers and classify customers in the high marketing responsiveness group if e-mails or catalog elasticities are greater than the respective median elasticities (for details, see Table 4, Panel B).

We used the aforementioned decision rules to classify each learner into one of the four decision process categories. We then “cross-tabbed” category membership by decision stage to examine the patterns by which customers evolve from one decision process category to another. Figure 3 shows these evolutions. The key findings are as expected: a tendency to evolve from non-preference-based to preference-based decision making and a tendency to evolve to lower marketing responsiveness. In particular, Figure 3 highlights two main types of evolutions: from Trial Category 3 to Posttrial Category 1 (42% of all learners); and from trial Category 3 to Posttrial Category 2 (51%). Both these evolutions are from a non-preference-based to a preference-based decision process.

Customers in Category 3 are responsive to marketing and non-preference based. This suggests that marketing bounces them from channel to channel during the trial stage. However, these customers eventually form strong channel preferences, which 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.

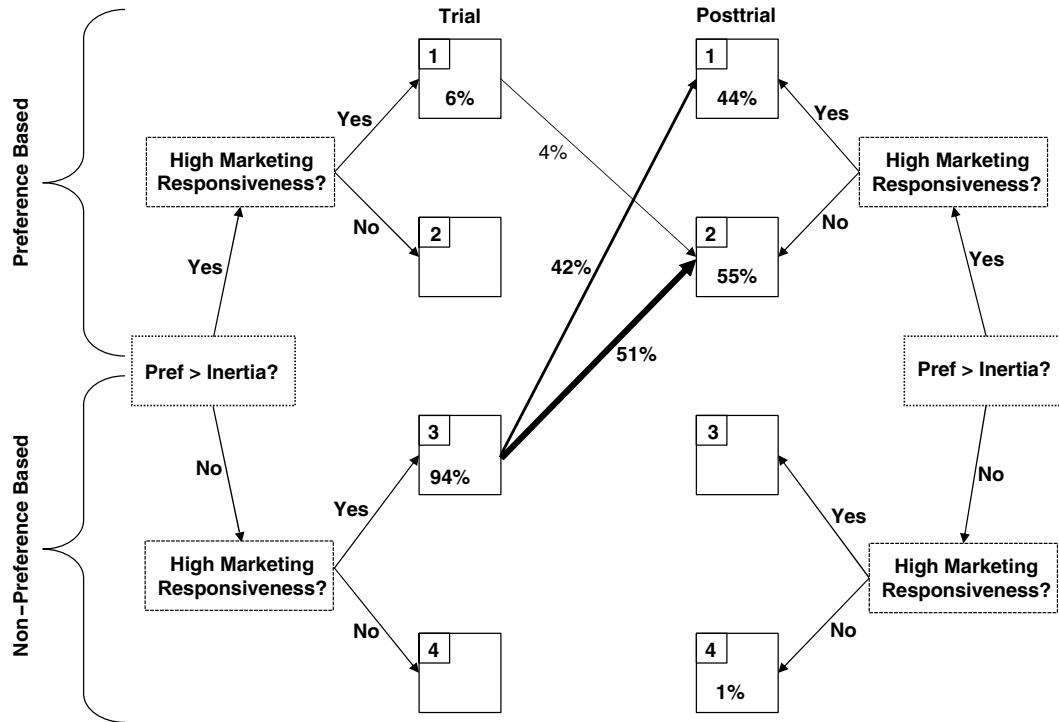
## Robustness Checks

### *Robustness with Respect to Assumptions*

We conducted robustness checks of some of our key assumptions by (1) investigating whether observed heterogeneity and time variance influence the customer’s propensity to evolve and (2) modeling a relationship between trial and posttrial parameters. To investigate the first issue, we estimated models with  $q_h$  a function of customer variables (age, gender: Model A) and time-varying variables including marketing (Model B); lagged channel choice, lagged product returns, age, and gender (Model C); and lagged product returns (Model D). We found that none of these models improved over our proposed model, which assumes that  $q_h$  varies randomly across customers and does not change over time (see Table 5). The proposed model’s superiority to Model B indicates that marketing does not significantly affect the length of the trial period. The comparison with Model D suggests that at least one measure of unsatisfying experience—product returns—does not encourage learning. A possible reason for this is that returning a book is not an unsatisfying experience. This is in line with previous work showing that when firms effectively manage service failure, a negative experience might turn into a positive one. Product returns can contribute to generate customer satisfaction and reinforce the relationship with the firm (Petersen and Kumar 2009).

To investigate the second issue, we posited a general correlation matrix in model parameters across both stages or made the posttrial parameters functions of the trial parameters plus age, gender, and first channel chosen. None of these models fit better than the proposed model, or they

**FIGURE 3**  
**Estimated Trial to Posttrial Learner Migration Patterns**



Notes: Only paths greater than 2% threshold are represented. From Table 5, Panel B, customers are classified as more 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) are 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. Although 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 posttrial stage. Using this rule, we find that 100% of the learners are highly marketing responsive in the trial period and 44% in the posttrial period. We also tested a more restrictive classification rule, requiring two marketing elasticities, not one, to be greater than the median. Using this rule, the number of responsive consumers decreases (from 100% to 91.6% in the trial period, and from 44% to 3.6% posttrial), but the same general finding holds: Learners are less marketing sensitive after they change.

did not converge (for details, see the Web Appendix W6 at <http://www.marketingpower.com/jmnov11>).

Overall, we find that the relatively simple proposed model performed better than plausible though more complex models. It might be that the more complex models overtaxed the data or that the proposed model captures the essence of decision process evolution, so that elaborations on that model do not add significantly more explanatory power.

### Replication in a Different Business Context

We perform a second empirical analysis to validate our results in a different business context. We obtained data from a U.S. retailer that sells durables and apparel through catalogs, stores, and the Internet. The company opened two stores in July 2000 and May 2001 that were relatively close to each other. We select a sample of customers who first purchased from one of these stores after May 2001 and lived within 15 miles of these stores. Therefore, we include only customers for whom the catalog, Internet, and store channels were available. The observation period is May 2001–September 2004. We include only active customers who made at least three purchases. The final sample size is 506 households. We have information on channel choice and catalog communications. Table 6 summarizes the data.

This is a useful data set because the opening of the stores is a discontinuity in the customer experience. Our theory posits that the decision process evolves when consumers have the motivation and ability to learn and are unfamiliar with the task. These conditions apply to our first data set, because the customer was just acquired. We believe these conditions would hold in the case of the second data set as well, because there has been an abrupt change in the channel environment. Indeed, Moe and Yang (2009) show that disruptive events force consumers to reexamine and possibly adjust their preferences and habits.

We compare our proposed model with the logit model that does not distinguish between trial and posttrial stages (Model 1). We also compare with a multinomial logit that assumes that each customer changes decision process after the same a priori defined number (15) of purchase occasions (Model 2).<sup>6</sup> The DIC statistics confirm the superiority of the proposed model over Models 1 and 2 ( $DIC_{\text{proposed}} = 11,711$ ,

<sup>6</sup>We use 15 because this was the median number of purchase occasions. We tested other numbers and also found the proposed model was superior ( $DIC = 12,085.5$  for 3 purchase occasions, 11,995.7 for 5, 12,371.5 for 12, and 12,275.5 for 18 versus  $DIC = 11,711.0$  for the proposed model).

**TABLE 5**  
**Robustness Checks for the Proposed Model**

Model	Description	DIC	Hit Rate
Proposed model	Two-stage multinomial channel choice model (Equations 2–5)	6525.8	72.9%
Model A	Two-stage multinomial channel choice model (with $q_{nt}$ a function of age and gender)	6598.2	53.1%
Model B	Two-stage multinomial channel choice model (with $q_{nt}$ a function of marketing)	6999.4	49.2%
Model C	Two-stage multinomial channel choice model (with $q_{nt}$ a function of age, gender, lagged number of channel used, and lagged returns)	7284.8	43.9%
Model D	Two-stage multinomial channel choice model (with $q_{nt}$ a function of lagged returns)	7202.8	42.4%

DIC<sub>model1</sub> = 12,082, DIC<sub>model2</sub> = 12,545). This provides reconfirmation of  $H_1$  and  $H_2$ , that the customer decision process evolves and that this evolution is heterogeneous across customers. The expected value of the probability of changing by the end of the data collection is .35, which means that we expect 35% of customers to have changed their decision processes. The average probability of changing for stayers is .25 and for learners is .53. We also find that the expected number of purchase occasions a learner needs to update his

**TABLE 6**  
**Descriptive Statistics for the Durable and Apparel Data Set (n = 506)**

Variable	M	SD	Min	Max
Average number of purchase occasions (per year)	5.7	3.1	1.0	20.0
Average number of purchase occasions over relationship	14.9	5.7	9.0	30.0
Average number of returns (per year)	.6	1.1	.0	10.0
Average number of catalogs received (per week)	.6	.8	.0	6.0
Age (years)	53.4	11.2	33.0	97.0
<b>Channel Usage<sup>a</sup></b>				
Mainly catalog	9.3%			
Mainly Internet	5.7%			
Mainly store	13.6%			
Catalog and store	20.6%			
Catalog and Internet	4.0%			
Internet and store	3.8%			
Catalog, Internet, and store	43.1%			

<sup>a</sup>“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.

or her decision process is 11. In contrast, it takes an average of 39 purchase occasions for a stayer to change.

Table 7 presents the estimates for the proposed model. The intercept estimates show that the store is preferred over both the catalog and the Internet in the trial stage, and the magnitude increases considerably in the posttrial stage. However the pseudo t-statistics for the preference parameters decrease in going from the trial to posttrial stage (from 7.7 to 2.1 for catalog vs. store; from 9.6 to 1.9 for Internet vs. store). The number of catalogs sent is significant in the trial stage but not in the posttrial stage. The pseudo t-statistics for catalogs sent decrease from the trial to posttrial stages (from 5.1 to 0.0 for catalog vs. store; from 2.5 to .05 for Internet vs. store), consistent with the changes in magnitude. In summary, we cannot clearly support an increase in the importance of preference going from the trial to post-trial stage. However, we find clear evidence for a decrease in the impact of marketing.

## Managerial Implications: Identifying Learners and Marketing to Them

Our results have established that there is a learner segment—a segment that significantly changes its decision process within a relatively short time. However, two questions naturally arise: In concrete terms, who is the learner segment? and Can we use this information to design more effective marketing programs? We answer the first question with a discriminant analysis and the second with a profitability scenario of a “right-channeling” strategy.

We hypothesize five variables that should distinguish between learners and stayers:

1. *Acquisition channel*: Our focal company predominantly uses “street agents” or “door-to-door” agents to acquire customers. Street agents approach customers directly outside the company’s bookstores; door-to-door agents go directly to prospects’ residences. We hypothesize that customers acquired through door-to-door agents would be less familiar with the company and its various marketing channels. Therefore, they would be more open to information and open to learning about the various channels (i.e., they would more likely be learners).
2. *Age*: Younger customers should be more open to exploring different channels and thus are more likely to be learners.
3. *Gender*: We have no particular hypotheses on gender, but given the literature on gender differences in shopping, we thought it was worthwhile to explore this variable.
4. *Immediate e-mail suppliers*: The firm asks for customers’ e-mails only when the customer is acquired. Some customers supply their e-mail addresses immediately, while others wait before doing so. Therefore, immediate e-mail suppliers are customers who provided their e-mail addresses at the time of acquisition. We hypothesize that the immediate suppliers would be more open to information and thus more likely to be learners.
5. *Big city dwellers*: Customers who live in big cities should be more familiar with various shopping alternative and thus less likely to be learners.

Table 8, Panel A, shows the means of these variables for learners and stayers. Three of our hypotheses are confirmed: Learners are more likely to be acquired through door-to-door agents, they are younger, and they are more likely to

**TABLE 7**  
**Estimates of Parameter Means: Two-Stage Channel Choice Model (Durable and Apparel Data Set)**

	Catalog Versus Store		Internet Versus Store	
	Parameter <sup>a</sup>	Elasticity <sup>b</sup>	Parameter <sup>a</sup>	Elasticity <sup>b</sup>
<b>Intercept</b>				
Trial	-1.46 (.19)	—	-2.87 (.30)	—
Posttrial	<b>-33.25</b> (15.90)	—	<b>-8.66</b> (4.60)	—
<b>Catalogs Sent</b>				
Trial	.51 (.10)	.22	.35 (.14)	.17
Posttrial	-.01 (2.82)	.04	-.07 (1.44)	.09
<b>Distance</b>				
Trial	.14 (.02)	—	.14 (.03)	—
Posttrial	1.63 (1.08)	—	.54 (.45)	—
<b>State Dependence</b>				
Trial	-.09 (.06)			
Posttrial	<b>2.35</b> (2.34)			

<sup>a</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>b</sup>We computed elasticities at the mean value of the continuous variables and the modal of the categorical variables.

Notes: Bold indicates that the 95% posterior interval for the parameter does not include zero. The standard deviation of the posterior distribution is shown in parentheses.

**TABLE 8**  
**Discriminant Analysis and Profitability Scenarios of Marketing to Learners**

<b>A: Mean Characteristics of Learners and Stayers</b>						
Variable	Learners	Stayers	p-Value for Difference			
Door-to-door agent acquisition	60.9%	49.7%	.003			
Age	41.2 years	44.5 years	.002			
Gender	70.2% female	72.4% female	.525			
Immediate e-mail suppliers	24.9%	2.7%	.000			
Big city dwellers	52.0%	54.4%	.533			
<b>B: Classification Matrix of Learners Versus Stayers</b>						
	Actual Group Membership					
	Stayer	Learner	Total			
<b>Predicted Group Membership (Number)</b>						
Stayer	708	80	788			
Learner	85	145	230			
<b>Predicted Group Membership (Conditional Probability of Correct Prediction)</b>						
Stayer	89.8%	10.2%	100%			
Learner	37.0%	63.0%	100%			
Hit Rate = (708 + 145)/1018 = 83.8%						
<b>C: Profitability Analysis: E-Mail-Early Versus E-Mail-Late Strategy<sup>a</sup></b>						
Strategy	Profits per Purchase Occasion			Web Use		
	First Period	Second Period	Entire Period	First Period	Second Period	Entire Period
E-mail early	\$16.81	\$14.22	\$15.52	88.4%	49.4%	68.9%
E-mail late	\$11.08	\$13.99	\$12.54	5.9%	46.2%	26.0%

<sup>a</sup>We used customers who had at least seven purchases. We noted the number of e-mails each customer received. Then we reallocated them to a 70-30 split for the e-mail-early strategy and a 30-70 split for the e-mail-late strategy. The e-mails were used on purchase occasions 1, 2, and 3 (early) and 5, 6, and 7 (late), with the exact number depending on the number of e-mails the customer received and the allocation dictated by the strategy being simulated.

supply their e-mail address immediately. Our next step is to estimate the discriminant function to use for classification. Results are shown in Table 8, Panel B.<sup>7</sup> The hit rate is 83.8%, significantly better than the random assignment standard of 65.5% (p-value = .000) ( $c_{\text{fair}}$  in Lehmann, Gupta, and Steckel 1998, p. 663). The model does not perform equally well in classifying learners: 37% of learners are misclassified. However, 63% of our predicted learners are actually learners, whereas only 10.2% of the predicted stayers turn out to be learners. In summary, the discriminant analysis shows we can identify the learners in concrete terms and classify customers as either learners or stayers.

By definition, the learner segment changes its decision process relatively quickly after acquisition. To illustrate the implications of this and how our discriminant analysis could be used in marketing to this group, we conduct a profitability analysis involving right-channeling learners. We assume the firm's profit margin is higher for Internet purchases (50%) versus catalog or store purchases (30%). Therefore, the company wants to encourage learners to use the Internet, and it knows that e-mails tend to do this (Table 3). The question is, Should those e-mails be used directly after customer acquisition, or later, after the customer has had a chance to get "settled"? Our results suggest the "e-mail-early" approach is better because learners are more receptive to marketing directly after they have been acquired.

The e-mail-early strategy allocates e-mails toward the purchase occasions immediately after the customer has been acquired. The e-mail-late strategy allocates e-mails toward the later purchase occasions. We use the discriminant function to identify learners and the parameter estimates obtained at the individual level to simulate channel choices under each strategy. Table 8, Panel C, summarizes the results and shows that, as we expected, the e-mail-early campaign generates approximately 24% more profits (\$15.52 – \$12.54/\$12.54) by encouraging earlier usage of the less costly Internet channel. This analysis shows that learners can be identified sensibly with concrete measures, and profits can be leveraged through appropriate marketing communication activities.

## Conclusions

### Key Results and Implications

The central questions of this study are as follows: (1) Do customers evolve to a different channel choice process over time? (2) If so, how many customers evolve? (3) How can we characterize the decision process of customers who evolve? and (4) What are the managerial implications for channel management?

To answer these questions, we characterized the channel choice decision process using the three main factors that previous researchers have used to study channel choice: channel preference, state dependence or inertia, and marketing. We hypothesized that (1) the channel choice

decision process evolves over time, (2) the time of this evolution varies across customers, (3) channel preferences become more relevant over time, (4) state dependence becomes less relevant over time, and (5) marketing communication is less effective over time. We developed a two-stage channel choice model to test these hypotheses and answer our key questions. We investigated customer channel choice behavior in a contractual setting in the book retail industry. As a robustness check, we analyzed a U.S. retailer of durables and apparel. Our key findings and managerial implications are as follows:

- Customers' channel choice decision processes evolve over time. This provides managers with the strategic insight that new-to-the-channel customers need to be treated differently from mature customers in terms of channel decisions. Operationally, it means that managers must take into account how recently the customer has been acquired in predicting channel choices.
- A significant portion of customers are expected to shift (ranging from 22% [newly acquired customers] to 35% [new channel introduction]). This suggests that the customer's speed of evolution can serve as a basis for segmentation: learners versus stayers.
- The predominant pattern of evolution is from a more marketing-responsive to a less marketing-responsive decision process. This means that the newly acquired customer or the customer who experiences a new channel introduction is more receptive to right-channeling than the mature customer. In general, mature customers will be less amenable to marketing efforts aimed at getting them to try new channels. This suggests that channel strategies for mature customers should ensure these customers are very satisfied with the service experience of the channel(s) they have chosen to use.
- Learners can be characterized through discriminant analysis, and the discriminant analysis can be used to identify the learner segment. In our case of the book retailer, we found that learners are likely to be acquired through door-to-door agents, immediately provide their e-mail addresses, and be younger. The result is a concrete simulation of how managers can identify learners and how they might market to them. Policy simulation can be used to examine the impact of various marketing strategies. We illustrate how an e-mail-early strategy can increase learners' profitability. Transitioning consumers early into the cheapest Internet channel increased the company's profits by approximately 24%.

The key hypotheses that (1) the customer decision process evolves, (2) the evolution is heterogeneous across customers, and (3) customers become less responsive to marketing over time were confirmed in two databases in very different situations (after the customer is acquired by a book retailer and after the introduction of a new purchase channel by a durables/apparel retailer). We confirmed our hypotheses that the process becomes more preference based and less inertial over time for the book retailer database but not for the durables/apparel database. The importance of preference in the trial stage in the durables/apparel database could be due to customers being highly familiar with purchase channels in this industry. The type of store the company introduced was similar to that of other U.S. competitors. Therefore, adding a store was not an entirely new experience for the customer. In terms of the learning theory that motivates decision process evolution, customers were relatively familiar with the task, which diminished learning.

The most important conceptual contribution of this work is that the customer's channel choice process evolves over

---

<sup>7</sup>Note that in addition to the variables in Table 8, Panel B, we added fixed effects for zip code to control more specifically for the location of the customer's residence. We included 256 of these fixed effects in the discriminant analysis.

time. We are encouraged regarding the generalizability of this finding in two databases. The underlying theoretical reason for evolution is customer learning, stimulated by three conditions: motivation and ability, lack of task familiarity, and dissatisfying experiences. To the extent these conditions are present in other decision processes (e.g., brand choice, price search, service usage), we would expect this evolution to hold beyond the domain of channel choice.

We do not advocate our model as the only way to model this evolution; the main point is that evolution should be considered 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 (Figure 3).

### **Limitations and Further Research**

There are limitations to this study that provide opportunities for further research. First, our work is based on what we can observe, namely, customers' purchase histories. A well-designed panel survey could enrich the characterization of stayers and learners and, of course, of their decision processes.

Second, we estimated different specifications of the proposed model that relax the assumption of independence between trial and posttrial stages. Although our results provide support for the independence assumption, the data we used might not be strong enough to pick this relationship up and/or the alternative modeling specifications might overtax the model with too many parameters. Additional research could investigate this issue further.

Third, the population of interest in this study is heavy users who account for high sales. This focuses attention on the best customers and avoids including customers active only during the trial stage. Work in the future might examine the evolution of the channel choice decision process for light users while accounting for customer churn. This would allow for a more complete analysis of customer retention and profitability.

Fourth, we focus on channel choice for the purchase decision and do not investigate search behavior (Verhoef, Neslin, and Vroomen 2007). Further research might explore the choice of channels for both search and purchase and the evolution of these two decision processes, but this would require the availability of both search and choice longitudinal data.

Fifth, we examined the impact of a single disruption in the environment on the ensuing transition from a trial to a posttrial decision process. A more general analysis would observe several disruptions. Related to this is the question of whether, in the face of a disruption, the trial model starts from scratch or is related to the previous decision

process the customer was using. To explore this, we examined our new store opening data. In that case, we had data for a subset ( $n = 115$  customers) who purchased before the opening of the stores. For these customers, we estimated a binomial logit model for choosing between the catalog and the Internet during the pre-store-opening period. We estimated individual-specific coefficients and correlated them with the individual-specific coefficient we estimated for the trial model. We found a correlation of .38 between the intercept terms representing preference for catalog versus Internet and a correlation of .43 between the catalog response coefficients representing impact on catalog choice rather than Internet. These correlations were both statistically significant at the .05 level and, in our view, promising given the small sample size and noise level in the data. These results also reinforce our conjecture that there was no change in the importance of preference associated with the opening of the store because of already established task familiarity. Notably, and consistent with Moe and Yang's view of environment changes as disruptions of customer habits, we found virtually no correlation ( $-.02$ ) between the state dependence coefficients. Overall, this points to the plausibility of multievent modeling in which trial decision processes build on the decision process that existed before the market disruption.

Sixth, we did not examine the specific content of e-mails and catalogs as it pertains to channel choice. Certainly, the catalog, for example, listed store locations and included the URL of the company's website. However, to our knowledge, there were no time-specific campaigns to encourage customers to use various channels. An examination of the potential for such campaigns would be interesting grounds for further research.

Seventh, a significant finding of our work is that marketing information affects channel choice in the trial stage but does not affect the timing of moving from the trial to the posttrial stage. This suggests that evolution is a customer-specific trait. Our robustness checks verify this result, but it is possible that in different contexts, marketing could influence not only channel choice but also the speed of evolution. It may be possible to design the information content of marketing to influence the timing of evolution. This would be supported, for example, by Petty and Cacioppo (1986, p. 126), who suggest that the ability to learn can depend on the message. We believe this is a fertile area for further research in the context of learning new decision processes, the domain of our work. Eighth, the focus of this study is on the customer channel choice decision process. Further research might also investigate whether other aspects of the customers' decisions, particularly purchase incidence and purchase quantity, evolve over time.

---

## **REFERENCES**

- Aaker, D.A. (1971), "The New-Trier Stochastic Model of Brand Choice," *Management Science*, 17 (5), 435–50.
- Alba, Joseph W. and Wesley J. Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (4), 411–54.

- Ansari, Asim, Carl F. Mela, and Scott A. Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research*, 45 (February), 60–76.
- Bettman, James R. and C. Whan Park (1980), "Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision Processes: A Protocol Analysis," *Journal of Consumer Research*, 7 (3), 234–48.

- Blattberg, Robert C., Byung-Do Kim, and Scott A. Neslin (2008), *Database Marketing: Analyzing and Managing Customers*. New York: Springer.
- Fader, Peter S., Bruce G.S. Hardie, and Chun-Yao Huang (2004), "A Dynamic Change-point Model for New Product Sales Forecasting," *Marketing Science*, 23 (1), 50–65.
- Gollwitzer, Peter M. and Ute Bayer (1999), "Deliberative Versus Implemental Mindsets in the Control of Action," in *Dual-Process Theories in Social Psychology*, Shelly Chaiken and Yaacov Trope, eds. New York: Guilford, 403–422.
- Gönül, Füsün F., Byung-Do Kim, and Mengze Shi (2000), "Mailing Smarter to Catalog Customers," *Journal of Interactive Marketing*, 14 (2), 2–16.
- Granbois, Donald H. (1977), "Shopping Behavior and Preferences," in *Selected Aspects of Consumer Behavior: A Summary from the Perspective of Different Disciplines*. Washington, DC: U.S. Printing Office, 259–98.
- Heilman, Carrie M., Douglas Bowman, and Gordon P. Wright (2000), "The Evolution of Brand Preferences and Choice Behaviors of Consumers New to a Market," *Journal of Marketing Research*, 37 (May), 139–55.
- Hoch, Stephen J. and John Deighton (1989), "Managing What Consumers Learn from Experience," *Journal of Marketing*, 53 (April), 1–20.
- Johnson, Eric J. and J. Edward Russo (1984), "Product Familiarity and Learning New Information," *Journal of Consumer Research*, 11 (1), 542–50.
- Knox, George A.H. (2006), "Modeling and Managing Customers in a Multichannel Setting," working paper, the Wharton School, University of Pennsylvania.
- Lee, Leonard, and Dan Ariely (2006), "Shopping Goals, Goal Concreteness, and Conditional Promotions," *Journal of Consumer Research*, 33 (1), 60–71.
- Lehmann, Donald R., Sunil Gupta, and Joel H. Steckel (1998), *Marketing Research*. Reading, MA: Addison-Wesley.
- LeSage, James P. and Manfred M. Fischer (2010), "Spatial Econometric Methods for Modeling Origin-Destination Flows," in *Handbook of Applied Spatial Analysis: Software Tools, Methods, and Applications*, Manfred M. Fischer and Arthur Getis, eds. Berlin: Springer-Verlag, 409–434.
- MacInnis, Deborah J., Christine Moorman, and Bernard J. Jaworski (1991), "Enhancing and Measuring Consumers Motivation, Opportunity, and Ability to Process Brand Information from Ads," *Journal of Marketing*, 55 (October), 32–53.
- Mattila, Anna S. (2003), "The Impact of Cognitive Inertia on Post-consumption Evaluation Processes," *Journal of the Academy of Marketing Science*, 31 (3), 287–99.
- Meyer, Robert J. and Arvind Sathi (1985), "A Multiattribute Model of Consumer Choice During Product Learning," *Marketing Science*, 4 (1), 41–61.
- Myers, Joe, Evan Van Metre, and Andrew Pickersgill (2004), "Steering Customers to the Right Channels," *The McKinsey Quarterly*, (October), (accessed July 28, 2011), [available at [http://www.mckinseyquarterly.com/Steering\\_customers\\_to\\_the\\_right\\_channels\\_1504](http://www.mckinseyquarterly.com/Steering_customers_to_the_right_channels_1504)].
- Moe, Wendy W. and Sha Yang (2009), "Inertial Disruption: The Impact of a New Competitive Entrant on Online Consumer Search," *Journal of Marketing*, 73 (January), 109–121.
- Narayanan, Sridhar, Manchanda, Puneet, and Pradeep K. Chintagunta (2005), "Temporal Differences in the Role of Marketing Communication in New Product Categories," *Journal of Marketing Research*, 42 (August), 278–90.
- Neslin, Scott A., Dhruv Grewal, Robert Leghorn, Venkatesh Shankar, Marije L. Teerling, Jacquelyn S. Thomas, and Peter C. Verhoef (2006), "Challenges and Opportunities in Multichannel Customer Management," *Journal of Service Research*, 9 (2), 95–112.
- and Venkatesh Shankar (2009), "Key Issues in Multichannel Customer Management: Current Knowledge and Future Directions," *Journal of Interactive Marketing*, 23 (1), 70–81.
- Noble, Steve, Amy Guggenheim Shenkan, and Christiana Shi (2009), "The Promise of Multichannel Retailing," *The McKinsey Quarterly: McKinsey on Marketing*, (October), (accessed July 28, 2011), [available at [http://www.mckinseyquarterly.com/Marketing/Digital\\_Marketing/The\\_promise\\_of\\_multichannel\\_retailing\\_2448](http://www.mckinseyquarterly.com/Marketing/Digital_Marketing/The_promise_of_multichannel_retailing_2448)].
- Payne, John W., James R. Bettman, and Eric J. Johnson (1992), "Behavioral Decision Research: A Constructive Processing Perspective," *Annual Review of Psychology*, 43, 87–131.
- Petersen, Andrew and V. Kumar (2009), "Are Product Returns a Necessary Evil? Antecedents and Consequences," *Journal of Marketing*, 73 (May), 35–51.
- Petty, Richard E. and John T. Cacioppo (1986), "The Elaboration Likelihood Model of Persuasion," in *Advances in Experimental Social Psychology*, Vol. 19, Leonard Berkowitz, ed. New York: Academic Press, 123–205.
- Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde (2002), "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society, Series B*, 64, 583–61.
- Thomas, Jacquelyn S. and Ursula Y. Sullivan (2005), "Managing Marketing Communications with Multichannel Customers," *Journal of Marketing*, 69 (October), 239–51.
- Tversky, Amos and Daniel Kahneman (1974), "Judgment Under Uncertainty: Heuristics and Biases," *Science*, 185 (4157), 1124–31.
- Venkatesan, Rajkumar, V. Kumar, and Nalini Ravishanker (2007), "Multichannel Shopping: Causes and Consequences," *Journal of Marketing*, 71 (April), 114–32.
- Verhoef, Peter C., Scott A. Neslin, and Björn Vroomen (2007), "Multichannel Customer Management: Understanding the Research-Shopper Phenomenon," *International Journal of Research in Marketing*, 24 (2), 129–48.
- Weiner, Bernard (1986), *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag.
- Weinberg, Bruce D., Parise, Salvatore, and Patricia J. Guinan (2007), "Multichannel Marketing: Mindset and Program Development," *Business Horizons*, 50, 385–94.

## **Web Appendix**

### **Decision Process Evolution in Customer Channel Choice**

Sara Valentini

Scott A. Neslin

Elisa Montaguti

#### **Table of Contents**

**APPENDIX W1** *Alternative Specifications to Estimate the Duration of the Trial Period*

**APPENDIX W2** *Estimation*

**APPENDIX W3** *Endogeneity Bias in Direct Marketing in Direct Marketing Communication Variables*

**APPENDIX W4** *Hit Rates and Cross-Validation*

**APPENDIX W5** *Calculation of the Expected Percentage of Learners*

**APPENDIX W6** *Contrasting Trial and Post-Trial Parameters*

**APPENDIX W7** *Dependence between Stages*

**APPENDIX W8** *Channel Switching Matrices*

**APPENDIX W9** *Model Free Evidence*

**APPENDIX W10** *Winbugs Program Used to Estimate the Model*



## APPENDIX W1

### ALTERNATIVE SPECIFICATIONS TO ESTIMATE THE DURATION OF THE TRIAL PERIOD

In our proposed model we estimate the length of the *trial* period using a geometric distribution, which assumes a “memoryless process” with a constant baseline hazard. This represents a parsimonious way to model the customer probability of changing from the *trial* to the *post-trial* stage. However, we also test alternative specifications using two different discrete distributions, that is the conditional Logit and the discrete Weibull.

#### *Conditional Logit Specification*

A conditional logit model that depends on the purchase occasion ( $t$ ) is given by:

$$(W1) \quad f(t) = \frac{1}{1 + e^{-(\nu_{0h} + \nu_{1h}t)}}, \quad t = 1, 2, \dots, T_h$$

Notice that  $\nu_{1h} = 0$  yields the geometric probability model whose constant hazard function indicates no duration dependence. If  $\nu_{1h} > 0$  the increasing hazard function indicates the longer the event the more likely it will happen. If  $\nu_{1h} < 0$  the decreasing hazard function indicates the longer the event the more likely it will not happen.

#### *Discrete Weibull Specification*

The Weibull distribution is given by:

$$(W2) \quad f(t) = (1 - q_h)^{t^{\beta_h}} - (1 - q_h)^{(t+1)^{\beta_h}}, \quad t = 1, 2, \dots, T_h$$

The baseline hazard can be flat ( $\beta=1$ ), monotonically increasing ( $\beta>1$ ), or monotonically decreasing ( $\beta<1$ ). This specification guarantees more flexibility over the geometric, which it nests as a special case (when  $\beta=1$ ).

We estimate these alternative specifications of the proposed model using a hierarchical Bayesian estimation in which we hypothesized that the "first stage" priors of the parameters  $v_{0h}$ ,  $v_{1h}$  follow a normal distribution with mean  $\mu$  and precision  $\tau$  and the "first stage" priors of the parameter  $\beta_h$  follow a gamma distribution. 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).

Table W1 summarizes the deviance information criterion (DIC) of the three alternative specifications. The DIC statistic confirms that the geometric specification outperforms the others suggesting that the baseline hazard does not depend on the number of purchase occasions elapsed.

**Table W1: DIC for Alternative Discrete Distribution Specifications**

	Geometric (proposed model)	Conditional Logistic	Discrete Weibull
DIC	6525.8	8070.4	7354.2

## APPENDIX W2 ESTIMATION

Heterogeneity in our parameters follows a normal distribution across customers with mean  $\mu$  and precision  $\tau$ . Our priors for these parameters follow a normal distribution for  $\mu$  and a gamma distribution for  $\tau$ . For  $\mu$ , we use the prior  $\mu \sim N(0,0.00001)$ , where 0.00001 is the precision. For  $\tau$ , we use the prior  $\tau \sim \text{Gamma}(0.5,0.5)$ . The mean of this prior is  $0.5/0.5 = 1$  and the variance is  $0.5/0.5^2 = 2$ . As a robustness check, we tried a gamma prior with parameters (3,1) and our main results were replicated. Note both of these are relatively informative priors. We obtained convergence problems when we used the highly “noninformative” priors (i.e.,  $\text{Gamma}(0.0001,0.0001)$ ) (see Gelman (Bayesian Analysis, 2006, No. 3, pp. 515-533) for problems when using this type of prior for heterogeneity in hierarchical models).

We use one million iterations for estimation and assess convergence for multiple chains using the Brooks-Gelman-Rubin convergence statistic and history graphs. Both suggest strong convergence see Figure W1 which shows the history graphs.

We verified that the estimation procedure was capable of recovering known model parameters through simulation assuming the mean parameters reported in Table 3 were the true parameters. We simulated different datasets by drawing individual parameters from a normal distribution with the means shown in Table 3. 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  $\alpha_{hj=1}^0 = .62$ ,  $\alpha_{hj=1}^1 = 1.25$ ,  $\alpha_{hj=2}^0 = -1.91^*$ ,  $\alpha_{hj=2}^1 = -2.22^*$ ,  $\beta_{1hj=1}^0 = -1.57^*$ ,  $\beta_{1hj=1}^1 = -.23$ ,  $\beta_{1hj=2}^0 = -7.55^*$ ,  $\beta_{1hj=2}^1 = -.05$ ,  $\beta_{2hj=1}^0 = .35$ ,  $\beta_{2hj=1}^1 = 1.28^*$ ,  $\beta_{2hj=2}^0 = 2.89^*$ ,  $\beta_{2hj=2}^1 = 0.33$ ,  $\beta_{3h}^0 = 2.80^*$ ,  $\beta_{3h}^1 = 1.17^*$ ,  $\mu_{c0} = -4.62^*$  (where  $*$  indicates that the parameter is significant at 5%). The correlation between the estimated coefficients in Table 3 and those recovered from the simulation was .98.

The coefficients that were statistically significant in Table 3 were also significant in the simulation and with the right signs. The model recovered the true parameters fairly well using the second dataset – the correlation between the coefficients in Table 7 and those recovered was .72. We believe that multicollinearity between catalogs and e-mails explains the lower, although still highly positive, correlation.

### APPENDIX W3

#### ENDOGENEITY BIAS IN DIRECT MARKETING COMMUNICATION VARIABLES

Both the catalogs sent ( $CS_{ht}$ ) and e-mails sent ( $ES_{ht}$ ) variables included in our channel choice model might be determined by firm's marketing strategy variables. For example, the European books retailer uses RFM metrics to determine the direct marketing communication policies (i.e. mailing decision). This means that the catalogs sent ( $CS_{ht}$ ) and e-mails sent ( $ES_{ht}$ ) variables in our channel choice model behave like endogenous variables, as they might be a function of the firm's direct marketing strategy and/or intrinsic customer characteristics.

Similarly to Gönül, Kim and Shi (2000), we use a two-stage least-squares approach to minimize this bias. This approach is in the same spirit as using instrumental variables instead of an endogenous variable on the right-hand side of an equation in a two-stage-least-squares framework. Specifically, our approach consists of two estimation stages. In the first stage, the endogenous variable becomes the dependent variable of a new regression model, where we use exogenous variables as covariates. Then, the predicted values of the dependent variable of this regression model are used in the second stage. This should minimize the problem of the correlation among the endogenous variables and the error terms.

Specifically, we estimate the probability of receiving a certain number of e-mails with a Poisson model. We use as explanatory variables seasonality dummies, lagged number of purchases and customer characteristics. Seasonality dummies indicate the quarters in which the company typically sends additional e-mails and catalogs (e.g. Christmas quarter or Mothers' Day). Lagged purchase number should control for the firm's RFM strategy. For example, more catalogs are sent to those customers who purchase more and more rapidly than others. Customer

characteristics (e.g. age, gender, Internet orientation) might account for differences in channel preferences.

We used the above described approach to handle endogeneity. First, we performed two regressions (an ordered logit for  $CS_{ht}$  and a poisson regression for  $ES_{ht}$ ). Second, instead of entering the actual values of catalogs sent and e-mails sent in the two-stage multinomial logit channel choice model we entered their predicted values. Table W2 and Table W3 display respectively the results of the e-mails and catalogs sent models for the books dataset. Table W4 shows the results of the catalogs sent model for the durable and apparel dataset. Please note that we do not have an e-mail sent model here because the apparel retailer did not provide us with data about potential e-mails campaigns of the company. Additionally, in this case we used weekly dummies, the amount (\$) of purchases made in each channel, the number of catalogs received before the first purchase, and a trend term as independent variables of the catalogs sent model.

The Pseudo  $R^2$  for the catalog models were a bit lower than desirable (.22, and .24). However, there were several significant variables in the equations and the predicted catalog values were strong enough to produce significant catalog effects in our choice model.

**Table W2: E-mails sent model parameter estimates (books dataset)<sup>a</sup>**

Poisson Regression		
Number of obs	=	14985
$\chi^2$ (7)	=	24447.8
Prob > $\chi^2$	=	.000
Max log-likelihood	=	-12025.6
Pseudo R <sup>2</sup>	=	.50
	Coefficient	Standard Error
Npo_lag	.11**	.02
Gender	.01	.40
Age	-.01**	.00
I_orient	3.54**	.05
Q <sub>1</sub>	-.03	.03
Q <sub>2</sub>	-.07	.03
Q <sub>3</sub>	-.21**	.03
Constant	-1.3**	.05
** significant at the .01 level		
* significant at the .05 level		

**Table W3: Catalogs sent model parameter estimates (books dataset)<sup>a</sup>**

Ordered logistic		
Number of obs	=	14985
$\chi^2$ (7)	=	4451.1
Prob > $\chi^2$	=	.000
Max log-likelihood	=	-7858.5
Pseudo R <sup>2</sup>	=	.22
	Coefficient	Standard Error
Npo_lag	1.09**	.03
Gender	-.05	.05
Age	.00*	.00
Q <sub>1</sub>	-3.76**	.08
Q <sub>2</sub>	-3.30**	.08
Q <sub>3</sub>	-3.44**	.08
cut1	-10.35	.46
cut2	-4.20	.10
cut3	1.81	.09
** significant at the .01 level		
* significant at the .05 level		
The store represents the baseline alternative		

**Table W4: Catalogs sent model parameter estimates (apparel dataset)<sup>a</sup>**

Poisson Regression <sup>b</sup>		
Number of obs	= 6661	
$\chi^2$ (175)	= 3222.6	Max log-likelihood = -5118.1
Prob > $\chi^2$	= .000	Pseudo R <sup>2</sup> = .24
	Coefficient	Standard Error
Catpurch	.01 <sup>**</sup>	.00
Netpurch	.01 <sup>**</sup>	.00
Storepurch	-.01 <sup>**</sup>	.00
Ncats_early	.02 <sup>*</sup>	.00
Trend	.04	.03
Constant	-.56 <sup>**</sup>	.18
<sup>**</sup> significant at the .01 level <sup>*</sup> significant at the .05 level <sup>b</sup> In this Table the coefficients of the 171 weekly dummies are not reported		

a

Q2= second quarter (Apr.-Jun.)

Q3= third quarter (Jul.-Sep.)

Q4= fourth quarter (Oct.-Dec.)

Npo\_lag= number of purchase occasions lagged.

Gender= 1 (female) 0 (male)

I\_orient= variable which assumes value 1 if the household has used at least once Internet as channel. It starts to assume value 1 after the first time that the internet choice happened.

catpurch= \$ total for catalog purchases

netpurch= \$ total for Internet purchases

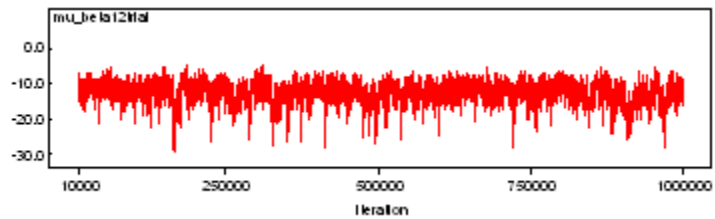
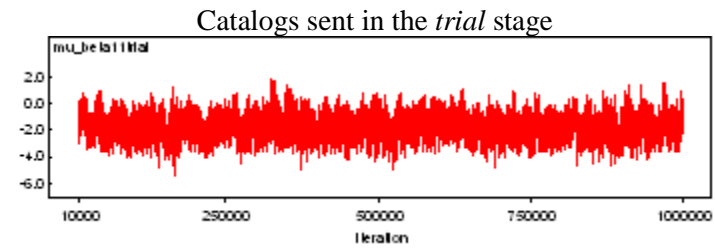
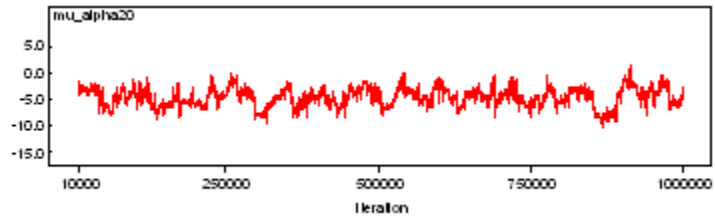
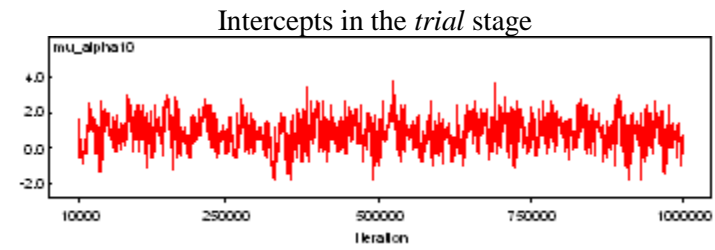
storepurch= \$ total for store purchases

Ncats\_early= # of catalogs received before the 4 weeks prior to the first purchase

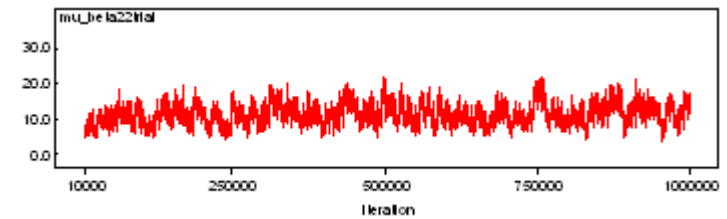
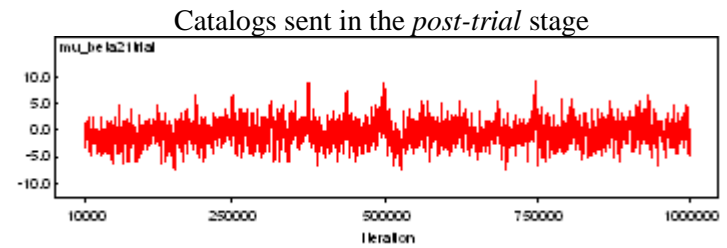
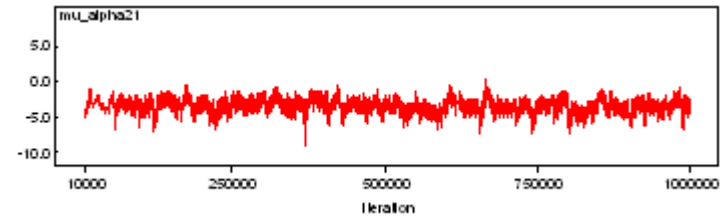
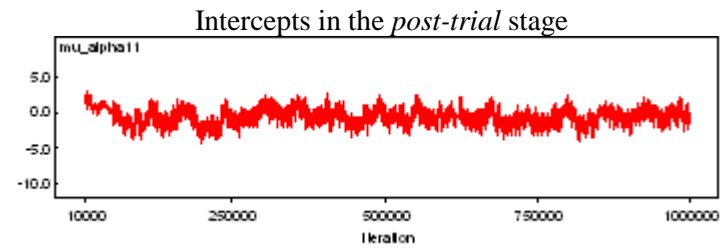
Trend= linear trend term



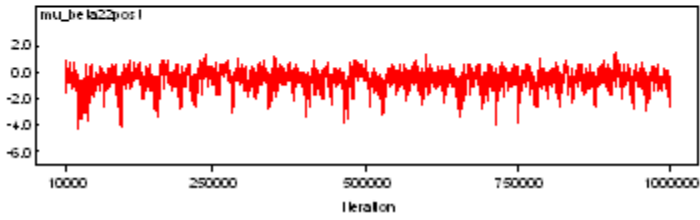
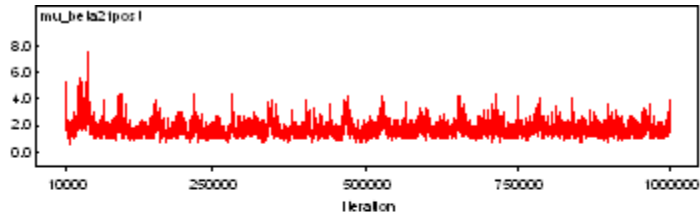
Figure W1<sup>a</sup>: History plots of the average posterior parameter estimates



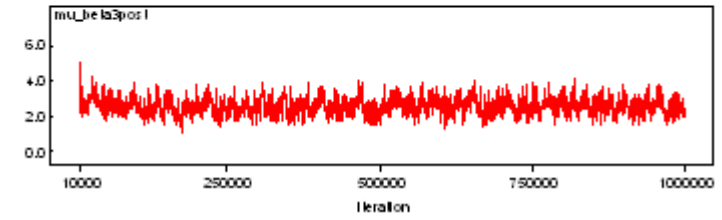
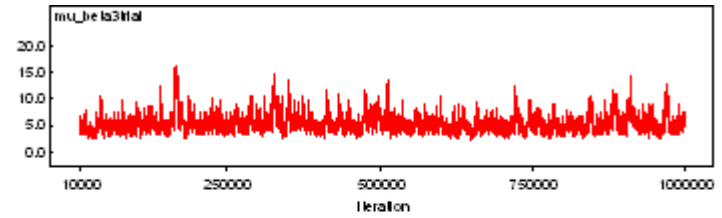
E-mails sent in the *trial* stage



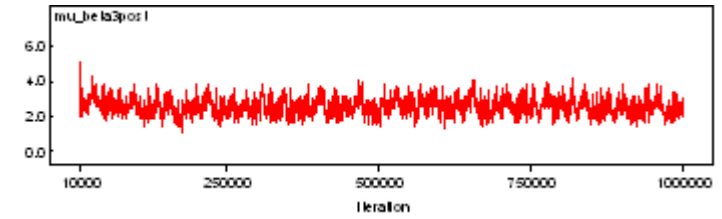
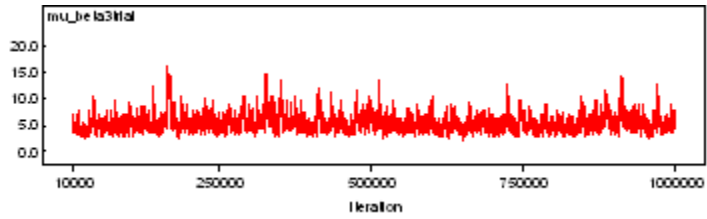
E-mails sent in the *post-trial* stage



State dependence - *trial* stage



State dependence - *post-trial* stage



<sup>a</sup> For illustrative purposes we reported only the history graphs we obtained similar results also for the Brooks-Gelman-Rubin convergence statistic.

*APPENDIX W4*  
*HIT RATES AND CROSS-VALIDATION*

We obtain each model's hit rate through cross-validation. Specifically we randomly select 75% (i.e. 770 customers and 11352 observations) for in-sample estimation and the remaining 25% (i.e. 248 customers and 3631 observations) for out-sample prediction. To predict the out-sample customers channel choices we proceeded in six steps: 1) We estimated the model using the in-sample data and we draw the average posterior estimates ( $\mu$ ) and standard deviations ( $\sigma$ ) from the joint distribution of  $\mu$  and  $\sigma$ . Specifically, we save the MCMC history of  $\mu$ 's and  $\sigma$ 's, 2) We drew samples of 248 values from the normal distribution specifying these  $\mu$  and  $\sigma$  as parameters of the generated variables. In this way, we simulated the individual level parameter for the out-sample customers, 3) We used such individual level simulated parameters to predict the channel choice probabilities of the out-sample customers, 4) We repeated steps 1-3 three times, 5) We computed the average choice probability for each customer and each purchase occasion, 6) We used these choice probabilities to assess the hit rate.

## APPENDIX W5

### *CALCULATION OF THE EXPECTED PERCENTAGE OF LEARNERS*

We compute the probability that customer  $h$  has evolved to the *post-trial* stage by her last observed purchase occasion ( $T_h$ ). To do this we substituted the estimated  $q_h$  for each customer, together with each customer's last purchase occasion ( $T_h$ ) into equation (2) (i.e., we used  $T_h$  for  $t$  in equation (2), where  $t=1, \dots, T_h$  for each customer  $h$ ). This was calculated for each iteration of the Bayesian estimation process and then averaged for each customer to obtain the probability the customer would have changed processes by the end of the observation period. The expected value of the probability of changing by the end of the data is .22, which means we expect 22% of customers to have changed decision process.

We believe 22% is the natural cut-off to use for defining a *learner* segment, i.e., those who are likely to change decision processes early, because this is the percentage of those who would be expected to change during our observation period. However, we conducted sensitivity analysis – using 30% and 40% as the cut-offs – and found that the key results in Figure 3 did not change, in fact there was even more of a tendency for *learners* to change to less marketing sensitive decision processes.

Similarly, we compute the number of purchase occasions customer  $h$  needs to evolve to the post-trial stage. Specifically, we substitute the estimated  $q_h$  for each customer, and the purchase occasion  $t$  into equation (2). In this way we obtain the probability that customer  $h$  will have changed process by purchase occasion  $t$ . For example, the expected value of the probability of changing process by the second purchase occasion is .04, which means that we expect 4% of customers to have changed decision process by the second purchase occasion. We ordered

customers according to their probability of changing by the second purchase occasion, and labeled the 4% highest in the second purchase occasion as those expected to change after 2 purchase occasions. We computed the expected value of the probability of changing decision process for each purchase occasion, obtaining the expected number of purchase occasions each customer needs to change decision process.

APPENDIX W6

CONTRASTING TRIAL AND POST-TRIAL PARAMETERS

We examined whether the *trial* versus *post-trial* differences in the coefficients of the proposed two stage channel choice model (reported in Table 3) were significant in the Bayesian sense, i.e., does the 95% credible interval for the differences exclude zero. We used 500K iterations to compute the posterior distribution for these differences and in all the cases reported in Table 3, the 95% credible interval excluded zero (see TableW5 for this test).

**Table W5: Bayesian Significance Test for Trial vs. Post-Trial Parameter Differences**

<b>Description</b>	<b>mean</b>	<b>2.50%</b>	<b>97.50%</b>	<b>SD</b>
$(\alpha_{hj=1}^0 - \alpha_{hj=1}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>intercept</b> coefficients (catalog vs. store)	.48	.43	.52	.02
$(\beta_{1hj=1}^0 - \beta_{1hj=1}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>catalogs sent</b> coefficients (catalog vs. store)	1.54	1.49	1.59	.03
$(\beta_{2hj=1}^0 - \beta_{2hj=1}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>emails sent</b> coefficients (catalog vs. store)	-1.54	-1.94	-1.14	.20
$(\alpha_{hj=2}^0 - \alpha_{hj=2}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>intercept</b> coefficients (internet vs. store)	-10.88	-11.13	-10.63	.13
$(\beta_{1hj=2}^0 - \beta_{1hj=2}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>catalogs sent</b> coefficients (internet vs. store)	-2.55	-2.62	-2.48	.04
$(\beta_{2hj=2}^0 - \beta_{2hj=2}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>emails sent</b> coefficients (internet vs. store)	2.88	2.79	2.96	.04
$(\beta_{3h}^0 - \beta_{3h}^1)$ =difference between <i>trial</i> and <i>post-trial</i> <b>state</b> <b>dependence</b> coefficients	1.01	.96	1.05	.02

## APPENDIX W7

### DEPENDENCE BETWEEN STAGES

The proposed model assumes independence between the *trial* and *post-trial* parameters. The underlying assumption of independence between customer's parameters across the two stages might generate biased estimates if the customer behaviors in the *trial* and *post-trial* period are related. For example, customer's preferences in the *trial* stage ( $\alpha^0_{hj}$ ) might be correlated to customer preferences in the *post-trial* period ( $\alpha^1_{hj}$ ). Similarly, customers who are receptive to catalogs in the *trial* ( $\beta^0_{1hj}$ ) might be also receptive to catalogs in the *post-trial* ( $\beta^1_{1hj}$ ).

We investigated this assumption by estimating two model specifications that allow for correlations between stages. In the first model we let the *post-trial* parameters (i.e.  $\alpha^1_{mhj}$ ,  $\beta^1_{mhj}$ ,  $\beta^1_{3h}$ ) be a function of the *trial* parameters (i.e.  $\alpha^0_{mhj}$ ,  $\beta^0_{mhj}$ ,  $\beta^0_{3h}$ ), demographic variables (age and gender) and an error term (Chernoff and Zacks 1964; Moe and Fader 2003). The second model added a dummy variable indicating the first channel selected as a covariate in the first model. We were unable to achieve convergence for the first model. The second model did achieve convergence but did not improve the fit of the model ( $DIC_{\text{proposed}}=6525.8$ ,  $DIC_{\text{model 2}}= 8014.0$ ). We then tried a more general formulation, allowing all the model parameters to follow a multivariate normal distribution. More specifically, we defined all the parameters in our model (both *trial* and *post-trial*) according to a multivariate normal distribution with a covariance matrix  $\Sigma$ .

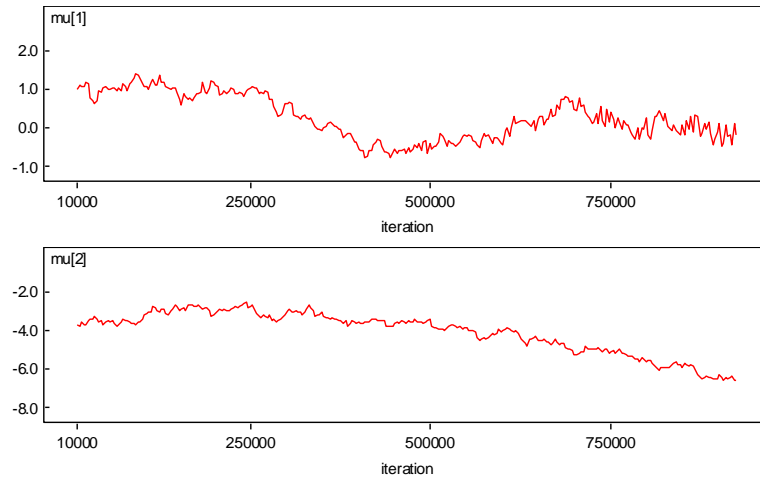
We use hierarchical Bayesian estimation for estimating this model. Specifically, we hypothesize that the priors of  $\alpha$ 's and  $\beta$ 's parameters follow a multivariate normal distribution with mean  $\mu$  and inverse covariance matrix  $\Sigma$ . The priors on the inverse covariance matrix follow a Wishart distribution. A screening of the iteration graphs of the posterior estimates

clearly highlighted that the model did not reach convergence after an extremely high number of iterations (i.e. a million). We then tried to help the model converge by judiciously selecting starting values. Specifically, we use the following procedure: (1) we run the model first without heterogeneity in the parameters; (2) we then used the estimates for the mean parameters from step 1 as the starting points of the independent-Normals-model (proposed model); (3) we used the estimates of the average posterior parameters and the individual-level parameters as the starting points of the model specification where all the logit function parameters are multivariate normal; (4) finally, we let this last model run for 1 million iterations. Figure W2 displays the history plots showing the average posterior estimates and highlights that this four step-procedure did not converged.

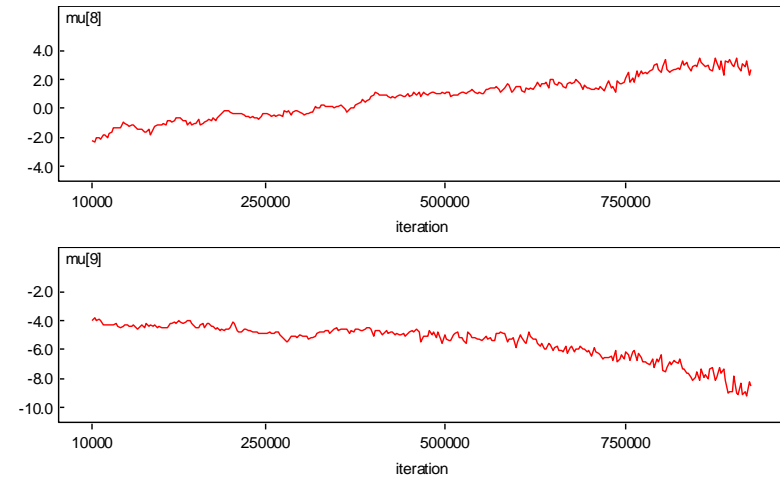


**Figure W2: History plots of the average posterior parameter estimates<sup>a b</sup>**

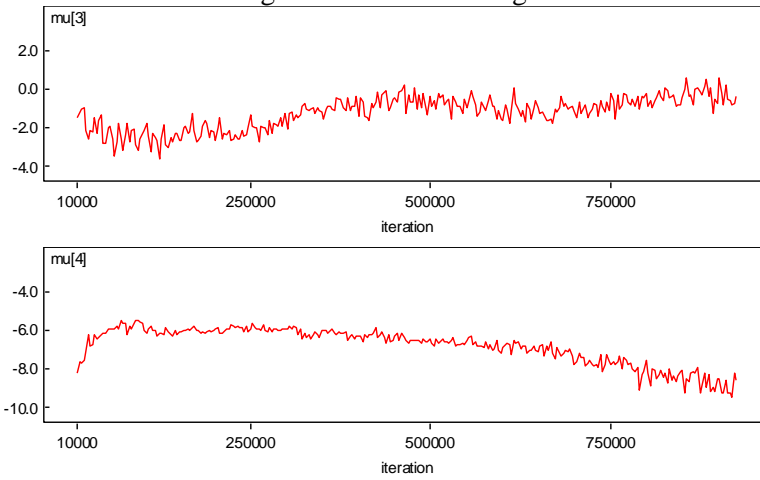
Intercepts in the *trial* stage



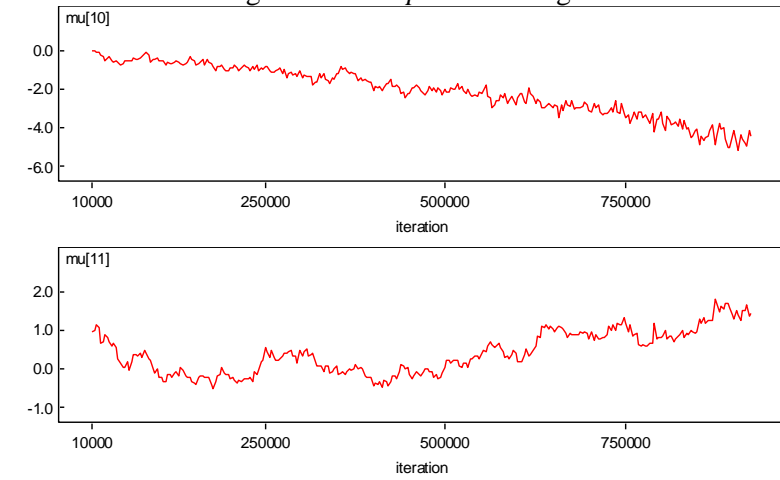
Intercepts in the *post-trial* stage



Catalogs sent in the *trial* stage

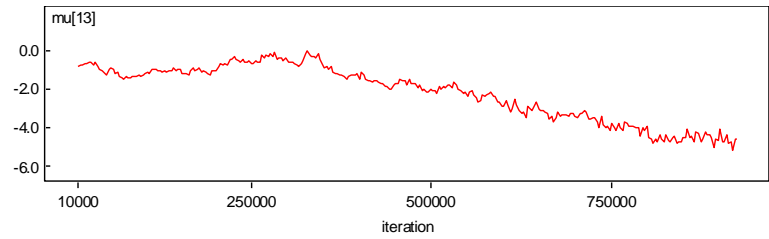
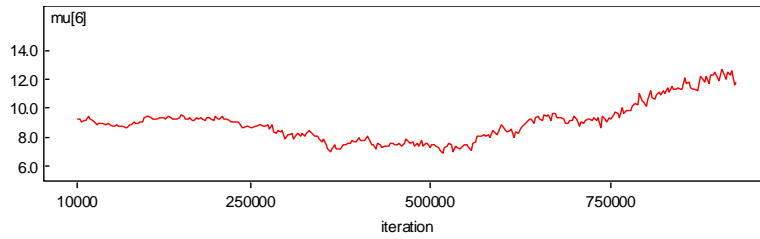
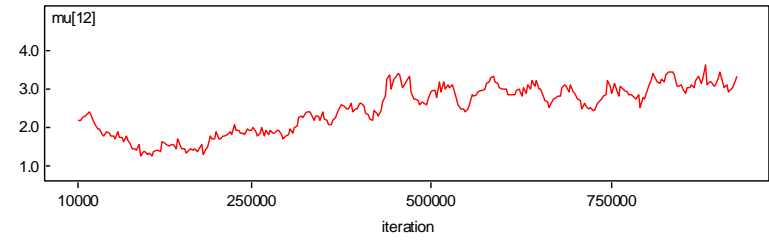
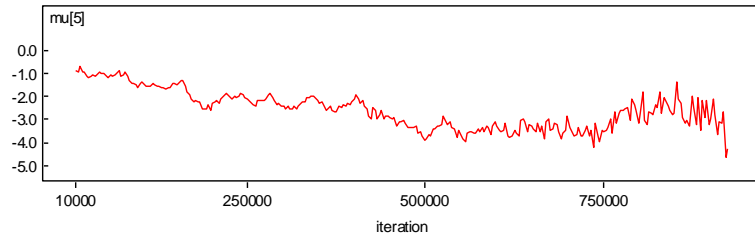


Catalogs sent in the *post-trial* stage



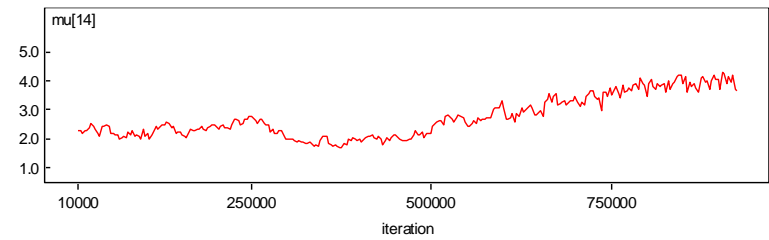
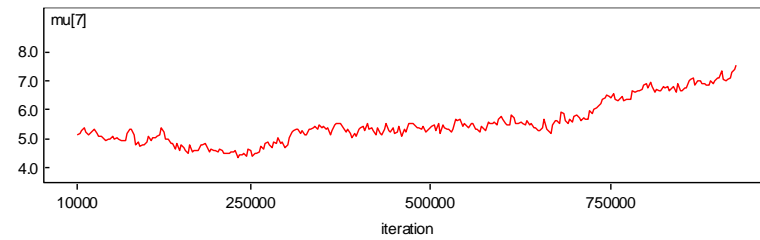
E-mails sent in the *trial* stage

E-mails sent in the *post-trial* stage

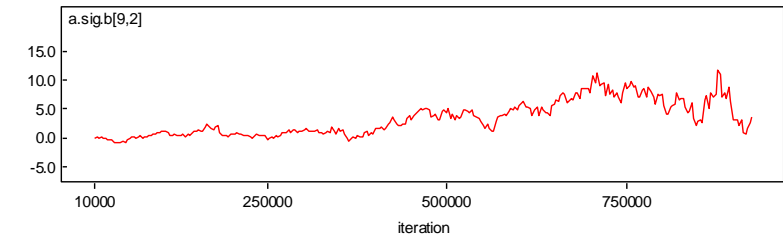
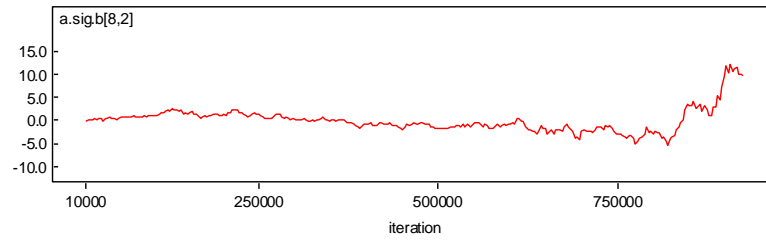
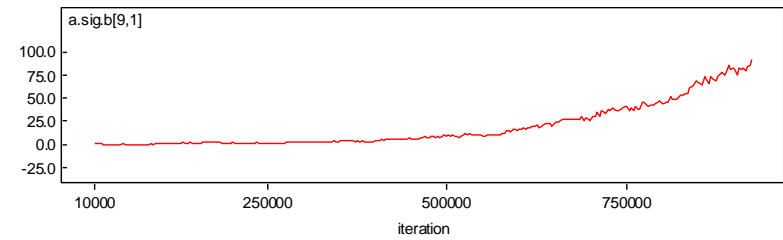
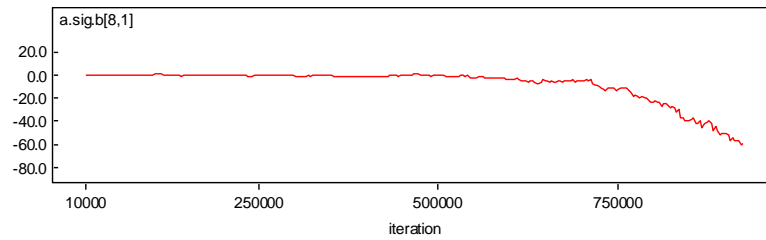


State dependence - *trial* stage

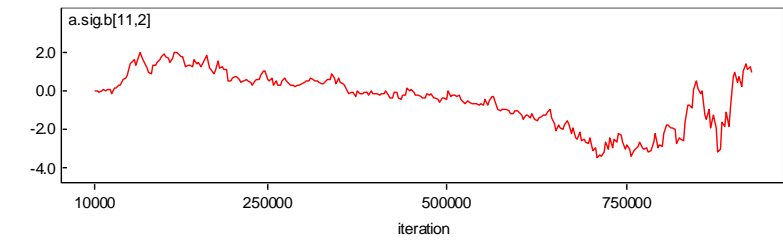
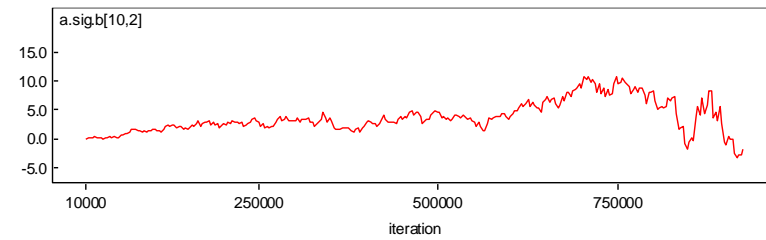
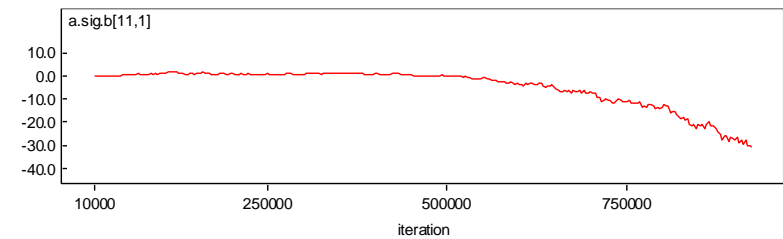
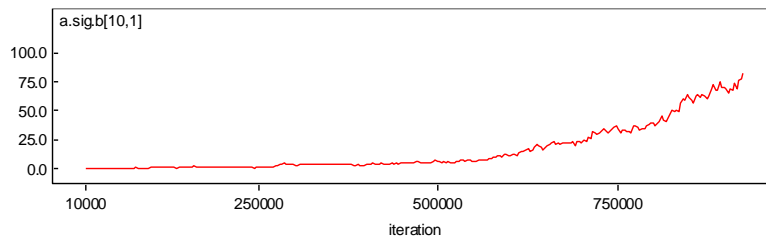
State dependence - *post-trial* stage



### Covariance parameters between channel preferences across *trial* and *post-trial*



### Covariance parameters preferences and marketing across *trial* and *post-trial*



<sup>a</sup> To get the plots we a sample of every 1/1000 iterations has been stored. This helped us to reduce storage requirement problems.

<sup>b</sup> For illustrative purposes we reported only the history plots representing the covariance matrix parameters which account for correlations between preferences and between preferences and marketing (both across *trial* and *post-trial*).

APPENDIX W8

CHANNEL SWITCHING MATRICES

**Table W6: Channel Switching matrix for the first period (absolute frequencies)**

FIRST HALF OF THE DATA (absolute frequencies)				
	Catalog	Internet	Store	TOTAL
Catalog	2658	65	82	2805
Internet	56	252	9	317
Store	83	22	3725	3830
TOTAL	2797	339	3816	6952

**Table W7: Channel Switching matrix for the first period (conditional switching probability)**

FIRST HALF OF THE DATA (conditional switching probability)				
	Catalog	Internet	Store	TOTAL
Catalog	94.8%	2.3%	2.9%	100%
Internet	17.7%	79.5%	2.8%	100%
Store	2.2%	.6%	97.3%	100%

**Table W8: Channel Switching matrix for the second period (absolute frequencies)**

SECOND HALF OF THE DATA (absolute frequencies)				
	Catalog	Internet	Store	TOTAL
Catalog	3079	142	106	3327
Internet	127	388	15	530
Store	155	17	4489	4661
TOTAL	3361	547	4610	8518

**Table W9: Channel Switching matrix for the second period (conditional switching probability)**

SECOND HALF OF THE DATA (conditional switching probability)				
	Catalog	Internet	Store	TOTAL
Catalog	92.5%	4.3%	3.2%	100%
Internet	24.0%	73.2%	2.8%	100%
Store	3.3%	.4%	96.3%	100%

## APPENDIX W9

### MODEL FREE EVIDENCE

We undertook several analyses to derive model-free evidence for two phenomena: (1) the existence itself of two different choice processes – *trial* and *post-trial*, and (2) the existence of inertial customers who are also marketing sensitive.

*Existence of two choice processes:* We performed the following model-free check to show that two different channel choice processes exist. Specifically, we took the first and second  $T_h/2$  observations for each customer. We estimated two very simple multinomial logit models (not hierarchical Bayesian estimation) which do not have heterogeneity and in which the change in the decision process is not modeled at all. The dependent variable is channel choice and the independent variables are the three main components of the decision process (i.e. state dependence, channel choice constants, and marketing). The first model considers only the first period (i.e. 1, ...,  $T_h/2$ ), by contrast the second model considers the second period (i.e.  $(T_h/2)+1$ , ...,  $T_h$ ). The analyses are based on simple logit model calculations where we do not model the evolution between these stages, but capture the phenomenon in a non-model based way – i.e., by dividing purchase occasions in half. So in that sense it is model-free evidence of evolution. Table W10 shows that the two models are different, implying the decision process changed. For example, the catalog intercepts become stronger in the second period, suggesting more preference-based decision-making, and the marketing variables, particularly catalogs, generally become smaller.

**Table W10: Model-free evidence: First and second period multinomial Logits**

Multinomial logit FIRST PERIOD			Multinomial logit SECOND PERIOD		
Number of obs	=	7718	Number of obs	=	7267
$\chi^2$ (8)	=	8719.4	$\chi^2$ (8)	=	11330.2
Prob > $\chi^2$	=	.000	Prob > $\chi^2$	=	.000
Max log-likelihood	=	-2216.4	Max log-likelihood	=	-946.4
	Coefficient	Standard Error		Coefficient	Standard Error
Intercepts			Intercepts		
Catalog	1.20**	.18	Catalog	-4.99**	.30
Internet	-.79*	.39	Internet	-6.42**	.65
N. Catalogs Sent			N. Catalogs Sent		
Catalog	-2.23**	.10	Catalog	.33**	.15
Internet	-2.78**	.27	Internet	-.59**	.28
N. Emails Sent			N. Emails Sent		
Catalog	.23**	.10	Catalog	.12	.08
Internet	1.05**	.08	Internet	1.22**	.10
State Dependence			State Dependence		
Catalog	8.36**	.21	Catalog	7.37**	.19
Internet	6.35**	.49	Internet	6.15**	.52
** significant at the .01 level					
* significant at the .05 level					
The store represents the baseline alternative					

*Existence of inertial customers who are marketing sensitive:* We took each customer and we defined him/her as an “inertial learner” if he/she made at least 2 contiguous (i.e.  $t-1$  and  $t$ ) choices of the same channel and then switched to a different channel. Then, we calculated the level of marketing at the time of the channel switch, compared to the average level when the customer was inertial. The t-test in Table W11 shows that higher marketing is associated with breaking the inertia of customers. Hence, inertial customers can respond to marketing, if there is enough of it.

**Table W11: Model-free evidence: Means comparison**

	Level of marketing (number of emails and catalogs received) when the “inertial learner” is inertial	Level of marketing (number of emails and catalogs received)for the “inertial learner” at the time of the channel switch
Mean	3.48	4.71
Std. Dev.	2.93	3.54
<i>t test</i> (mean comparison)	Ha: diff > 0 $t = 10.07$ $df = 3852$ $Pr(T > t) = .000$	

Additionally, we classified each household in our sample as inertial or non-inertial. We classify a customer as inertial if he/she made at least 2 contiguous (i.e.  $t-1$  and  $t$ ) choices of the same channel and then switched to a different channel or if he/she is a single channel user. We identified 722 inertial customers. Then we ran a simple multinomial logit model with channel choice as dependent variable and marketing as covariate only for the inertial customer group. Results are summarized in the Table W12, which shows that inertial people can respond to marketing, and in fact they do. This is in line with what our model finds in Figure 3.

**Table W12: Model Free Evidence: Multinomial logit on the Inertial Group**

Multinomial logit for the “Inertial” group (n=722)		
Number of obs	= 10835	
$\chi^2$ (4)	= 1214.4	
Prob > $\chi^2$	= .000	
Max log-likelihood	= -7067.7	
	Coefficient	Standard Error
Intercepts		
Catalog	-2.15**	.09
Internet	-2.29**	.93
N.Catalogs Sent		
Catalog	.87**	.49
Internet	-2.95**	.68
N.Emails Sent		
Catalog	1.01**	.32
Internet	2.88**	.33
** significant at the .01 level		
* significant at the .05 level		
The store represents the baseline alternative		

APPENDIX W10

WINBUGS PROGRAM USED TO ESTIMATE THE MODEL

```
model
{
for (i in 1:NHH){
q[i]<- max(0.0000, min(0.98, 1/(1+exp(-(c0[i])))))
for (s in 1:Tpo[i]) {
X1[i,s] <- 1-(pow((1-q[i]),(s-1)))
}
}

for (n in 1:NOBS) {

U0[n,1] <- b00[h[n],1]+b10[h[n],1]*CS[n]+b20[h[n],1]*ES[n] + b30[h[n]]*LC[n,1]
U0[n,2] <- b00[h[n],2]+b10[h[n],2]*CS[n]+b20[h[n],2]*ES[n] + b30[h[n]]*LC[n,2]
U0[n,3] <- b30[h[n]]*LC[n,3]

U1[n,1] <- b01[h[n],1]+b11[h[n],1]*CS[n]+b21[h[n],1]*ES[n] + b31[h[n]]*LC[n,1]
U1[n,2] <- b01[h[n],2]+b11[h[n],2]*CS[n]+b21[h[n],2]*ES[n] + b31[h[n]]*LC[n,2]
U1[n,3] <- b31[h[n]]*LC[n,3]

vbot0[n]<-exp(U0[n,1])+exp(U0[n,2])+exp(U0[n,3])
vbot1[n]<-exp(U1[n,1])+exp(U1[n,2])+exp(U1[n,3])

pr1[n,1]<- (1-X1[h[n],po[n]])*((exp(U0[n,1])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,1])/vbot1[n]))
pr[n,1]<-max(.00000000,min(.98, pr1[n,1]))

pr1[n,2]<- (1-X1[h[n],po[n]])*((exp(U0[n,2])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,2])/vbot1[n]))
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))

pr1[n,3]<- (1-X1[h[n],po[n]])*((exp(U0[n,3])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,3])/vbot1[n]))
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))

Y[n,1:3] ~ dmulti( pr[n,1:3] , 1)

}
```



```

# first stage PRIORS
for (n in 1:NHH) {

b30[n] ~ dnorm(mu30,prec30)
b31[n] ~ dnorm(mu31,prec31)

c0[n] ~ dnorm(muc0,precc0)

for (k in 1:2) {
  b00[n,k] ~ dnorm(mu00[k],prec00)
  b10[n,k] ~ dnorm(mu10[k],prec10)
  b20[n,k] ~ dnorm(mu20[k],prec20)
  b01[n,k] ~ dnorm(mu01[k],prec01)
  b11[n,k] ~ dnorm(mu11[k],prec11)
  b21[n,k] ~ dnorm(mu21[k],prec21)
}
}

# Choice model second stage priors

for (k in 1:2) {
mu00[k]~dnorm(0,.00001)
mu10[k]~dnorm(0,.00001)
mu20[k]~dnorm(0,.00001)
mu01[k]~dnorm(0,.00001)
mu11[k]~dnorm(0,.00001)
mu21[k]~dnorm(0,.00001)
}
mu30~dnorm(0,.00001)
mu31~dnorm(0,.00001)

prec00~dgamma(0.5,0.5)
prec01~dgamma(0.5,0.5)
prec10~dgamma(0.5,0.5)
prec11~dgamma(0.5,0.5)
prec20~dgamma(0.5,0.5)
prec21~dgamma(0.5,0.5)
prec30~dgamma(0.5,0.5)
prec31~dgamma(0.5,0.5)

# geometric second stage priors
muc0~dnorm(0,.00001)
precc0~dgamma(0.5,0.5)

```

}

## References

- Chernoff, Herman and Shelley Zacks (1964), "Estimating the Current Mean of a Normal Distribution Which Is Subject to Changes in Time," *Annals of Mathematical Statistics*, 35 (3), 999–1018.
- Gönül, Füsün F., Byung-Do Kim, and Mengze Shi (2000), "Mailing Smarter to Catalog Customers," *Journal of Interactive Marketing*, 14 (2), 2–16.
- Moe, Wendy W. and Peter S. Fader. 2003, "Capturing Evolving Visit Behavior in Clickstream Data," *Journal of Interactive Marketing*, 18 (Winter) 5–19.

Copyright of Journal of Marketing is the property of American Marketing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.