The diffusion of a new service: Combining service consideration and brand choice

Vardit Landsman · Moshe Givon

Abstract We propose an individual-level model of a two-stage service diffusion process. In the first stage, customers decide whether to "consider" joining the service. This (Consideration) stage is modeled by a hazard model. Customers who decide to consider the service move on to the Choice stage, wherein they choose among the service alternatives and an outside No Choice option. This stage is modeled by a conditional Multinomial Logit model. The service provider does not observe the transition in the first stage of potential customers who have yet to choose a brand. Such potential customers may have started to consider joining the service, yet chose the outside alternative in each period thereafter. One of the main contributions of the model is its ability to distinguish between these two non-adopter types. We estimated the model using data on the adoption process of newly introduced service plans offered by a commercial bank. We employed the hierarchical Bayes Monte Carlo Markov Chain procedure to estimate individual as well as population parameters. The empirical results indicate that the model outperforms competing models in breadth of analysis, model fit, and prediction accuracy.

Keywords Diffusions · Brand choice · Bayesian analysis

JEL Classification C5 · M3 · D9 · L8 · O3

1 Introduction

The adoption of an innovation, from the perspective of potential customers, may involve two major processes. First, a potential customer may undergo various pre-

V. Landsman (✉)
Erasmus School of Economics, Erasmus University, Rotterdam, the Netherlands
e-mail: landsman@few.eur.nl

M. Givon
Recanati School of Management, Tel -Aviv University, Ramat Aviv, Tel Aviv 69978, Israel
e-mail: givon@post.tau.ac.il

adoption states, such as Awareness or Consideration, before finally adopting the new service / product. Second, actual adoption requires an active choice process among available alternatives in the category. Consequently, the non-adopter population is likely to be heterogeneous with respect to its stage in the adoption process. This heterogeneity, while having important marketing implications, is not captured by typical innovation diffusion models.

The following example may illustrate this point. Consider the introduction of the online DVD rental service to the U.S. market during the late 1990s. This service offered mail delivery of disks from a predefined list of preferred DVDs to rent. Netflix, the prototype for the industry, was soon followed by competitors such as Blockbuster, RedBox, and even Wal-Mart. Membership plans for the service differ between and within service providers. The differences, for example, are in the number of disks sent at fixed intervals (often weekly), or the ability to rent individual movies for a fixed fee with no monthly payment.

Let's focus on a hypothetical potential customer for this service. For a while after the service introduction, this customer may not even consider joining it for any number of reasons. She might, for instance, not be aware of the new service, or perceive no obvious benefit from it. At some point, she might start considering the service, again for many possible reasons, such as a recommendation from a friend or advertising. At this point, she begins to gather information about available service alternatives, i.e., competing membership plans. One possible outcome of this information search may be that none of the offers is found suitable, in which case this potential customer may "choose not to choose" and wait for something to change, such as a better offer.

From the point of view of a typical diffusion model, this potential customer's state has not changed since the introduction of the service. Yet this point of view fails to reflect the process the customer underwent. Typical diffusion models therefore cannot distinguish between potential customers who would not consider using the service and those that begin considering the service, do not find a suitable alternative, yet are now "in the market" for one. Insights concerning the sources of this heterogeneity, together with knowledge of the relative sizes of various potential adopter groups, can improve our understanding of the diffusion process and consequently our ability to manage it successfully.

In this study, we address this challenge and formalize an individual-level model of service diffusion and brand choice that takes into account the flow of potential customers between pre-adoption states. Taking such an individual customer and brand level perspective is consistent with recent calls by marketing scholars to further develop individual level and brand level viewpoints in the analysis of diffusion processes (Chatterjee et al. 2000; Muller et al. 2009; Roberts and Lattin 2000).

Potential customers in this model are not viewed as one homogeneous pool of potential market, as in a typical product category diffusion model. Instead, we can now distinguish between two non-adopter types: potential customers who have not even considered joining the service and potential customers, who have, at some point, started considering the service, yet concluded that none of the available service options suits their needs. Moreover, we also allow for an association between the influences of the factors affecting potential customers' inclination to consider the service and those affecting the choice between service alternatives (Chib et al. 2004).

The ability to distinguish between different types of non-adopters can serve as a diagnostic tool that depicts not only to what extent a new service is considered among potential customers, but also the differential effect that marketing variables have on service consideration and brand choice. In terms of managerial policy-making, the model can have strong relevance, as it may affect the way service providers seek to accelerate service adoption. There can be a big difference between the tools and strategies required to stimulate consideration of the new service, and those that aim to improve choice probabilities, and in particular to lower the number of considering potential customers who defer choice. Moreover, individual-level estimation of the model allows us to calculate indications as to the state of every potential customer. These, in turn, enable a targeted implementation of chosen marketing activities.

By knowing the likelihood that an individual belongs to each pre-adoption type, the firm can better define its objectives in direct marketing efforts. More specifically, while marketing communications towards the first non adopter type (i.e., the 'no consideration' type) should involve getting into the consideration set of the customer, those towards the second type of pre-adopters should aim at "closing the deal" either by providing new alternatives to the choice set or by trying to persuade the customer to reconsider his/her choice and provide arguments that will enable the choice of an available alternative.

The model that we propose is suitable mainly for continuous services (Bolton 1998) that are primarily subscription-based (Reinartz and Kumar 2000). Such services typically have an explicit start date, yet no fixed end date. There can be two main types of continuous services: new service categories (e.g., cable TV); and new services offered on top of an existing service (e.g., Video on Demand plans offered by a cable TV provider). The latter types of services are typically "service augmentations" that are built on top of a core service to differentiate it from competitors and increase customer benefits (Berry 2002). Given the economic growth in services, their profits and competitive advantage potential, and the overall decline in customer satisfaction with services (Zeithaml et al. 2006), it is of great importance to develop tools, such as the model suggested herein, that improve the ability to market and manage services.

Although the diffusion literature has generally modeled the diffusion of a new service in the same manner as that of a durable good (Hogan et al. 2003; Jain et al. 1991; Krishnan et al. 2000; Lilien et al. 2000) there are some distinct differences between services and durable goods that are essential for the modeling of a diffusion process, both from the theoretical and practical perspective. For instance, unlike the adoption process of durable goods, adopters of continuous services remain "involved" in the category in the sense that they can actively switch between service alternatives or quit the category altogether in any later post-adoption time period. This characteristic is taken into account in our model and provides, in addition to a better behavioral representation, additional insights for customer relationship management.

For the empirical implementation of the model, we use a unique data set portraying the adoption of a new service offered by a commercial bank to its active customers. The estimation results demonstrate the model's ability to provide insightful information on the adoption process as a whole, as well as on various

factors that affect the two consecutive processes of consideration and choice. We furthermore display how our "full" model outperforms competing models not only in analysis breadth, but also based on various information criteria and holdout sample prediction accuracy.

2 The model

Consider the diffusion of a new service in a discrete time setting that begins at service introduction. At this point in time, all potential customers are in the first of the three model states to which we refer as *No-Service No-Consideration* (see Fig. 1). Starting from the first time period, each potential customer has a non-negative probability of exiting this state and starting to consider joining the new service.[1] Once a potential customer exits the No-Service No-Consideration state, she embarks on a choice process between available brands and an outside option representing a decision to defer choice. In the latter case, the potential customer enters the second model state called *No-Choice*. Considering potential customers who do choose one of the service alternatives are in the third possible state denoted as *Choice*. It is important to note that Consideration is not a state of the model, but rather a beginning of a process that leads through the choice process to either the No-choice state or the Choice state, comprised of the service alternatives.

Looking at the model so far, we see that the process can be decomposed into two stages. The first is a dynamic, time-dependent stage representing the transition of potential customers from a state of Non-Consideration to Consideration. The second stage is a choice process that is not affected by the passage of time itself.

Potential customers who enter the No-Choice state can, at any later time period, choose an alternative and join the service. This transition is likely to be triggered by an introduction of a new alternative into the choice set, or by a change in any of the existing alternatives. In addition, customers of the service can, at any time period after adoption, decide to switch between alternatives, or quit the service all together. Quitting customers are modeled to enter the No-Choice state, and therefore can also, in any later time period, decide to rejoin the service. The fact that in our model, customers can move in and out of the service and switch between service alternatives at no cost at any time precludes the need for forward looking on the part of the customers.

---

[1] The term "consideration" in our setting is used to describe the point in time at which a potential customer starts evaluating the available alternatives for the service with the intention of choosing the most suitable alternative, or deferring choice if none of the alternatives is found suitable. This term has been used before in the context of diffusion by Weerahandi and Dalal (1992) to reflect a change in potential customers' attitudes toward the innovation. Specifically, the term was used to represent the point at which "the customer contemplates acquisition of the service". In their model, Weerahandi and Dalal regarded the consideration stage as a "diffusion stage" and used the Bass model to capture consideration probabilities (Bass 1969).
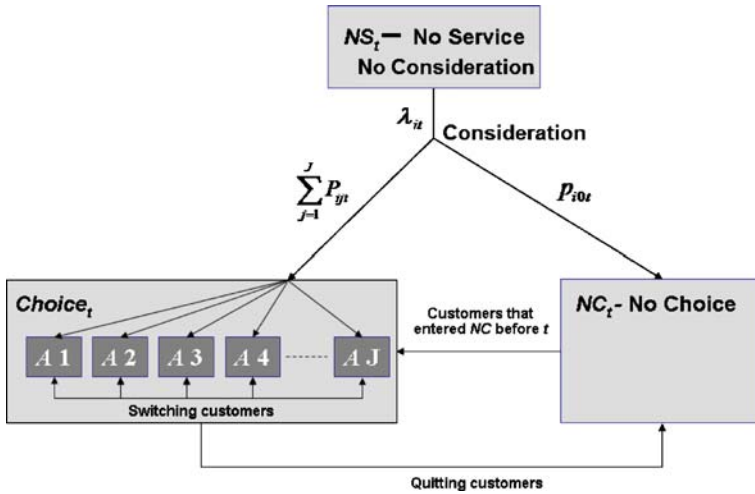
Fig. 1 State flow in the diffusion process

3 Relevant literature

The framework we propose integrates literature from two sub-disciplines in marketing: diffusion and choice. Accordingly, the model comprises elements from the corresponding diffusion and choice models.

3.1 Diffusion models

In the diffusion literature, we typically see dynamic models that predict category adoption as influenced by internal and external influences. The diffusion process in these models is viewed as a single-stage, binary-state process where at any point in time, individuals are either adopters or non-adopters. Moreover, typical diffusion models focus on category adoption rather than brand adoption.

There are exceptions wherein diffusion studies do take into account some aspects of our proposed model. There are, for instance, a few multi-state macro flow diffusion models that view the adoption process as gradual, and thus account for heterogeneity in customers' states by asserting that non-adopters go through Non-Awareness and Awareness stages (Dodson and Muller 1978; Mahajan et al. 1984; Kalish 1985); or Non-Consideration and Consideration stages (Weerahandi and Dalal 1992). However, this heterogeneity is not at the individual potential customer level, but rather at the aggregate level. Moreover, none of the models regard the diffusion process at the brand level.

There are also several diffusion models that account for multiple brands in the market (Parker and Gatignon 1994; Givon et al. 1995; Givon et al. 1997). All these models focus mainly on the aggregate effects of interpersonal influences and competition on the diffusion of multiple competing brands in a new category. Unlike the model proposed in this study, they do not investigate the diffusion in a choice setting, wherein the final choice is affected by brand and decision-maker attributes and tastes.

Another relevant study in the context of brand-level diffusion modeling is that of Krishnan et al. (2000; hereafter, KBK). The framework proposed in this study also attempts to model sales growth at the brand level and addresses the issues of potential market for category brands and its sources of influence. A central question raised in this framework is whether, in the product category, the "which brand to buy?" decision is secondary to deciding "whether or not to buy the category". In this sense, KBK addressed one of the major behavioral issues in our model, namely, the two stages of *Consideration* and *Choice*. However, unlike our basic assumption, KBK also considered a reverse order for the two questions. For example, they posited that in some categories, such as sports cars, the brand question might arise during initial stages of the adoption decision. Our proposed model focuses primarily on services for which category consideration comes before brand choice; this focus dictates the primary assumption guiding our model.

Another interesting and relevant branch of diffusion literature derives adoption behavior from individual-level utility maximization (Roberts and Urban 1988; Chatterjee and Eliashberg 1990; Horsky 1990). These studies, however, either do not focus on the sources of consumer state-heterogeneity as in our proposed model (e.g., Chatterjee and Eliashberg 1990; Horsky 1990), or, as in Roberts and Urban (1988), suggest a dynamic brand choice model that investigates the introduction of a new brand in an existing category, thus applicable mainly to categories where replacements are the dominant source of category sales.

A possible alternative to modeling time to adoption at the micro level is through an optimal stopping problem as suggested by Song and Chintagunta (2003). According to this approach, potential customers who adopt the product exit the market, i.e., they stop being active in the adoption process. Potential customers who opt not to adopt at time t remain active in the next time period, and therefore must make another decision about adoption at $t+1$. Song and Chintagunta implemented the optimal stopping problem approach to develop a consumer choice model for new product adoption that allows for consumer heterogeneity as well as consumers' forward-looking behavior. Their model is aimed mainly at durable goods, particularly high-technology product markets, wherein repeat purchase does not provide a significant source of sales, and price tends to decline over time while quality improves (as new features are typically introduced at a rapid pace). In such markets, forward-looking potential customers may anticipate this price and quality pattern and optimize the timing of their purchase by trading their utility from buying the product right away for expected future utility from differing price and quality levels.

The main difference between our proposed model and that of Song and Chintagunta (2003) lies in the nature of the innovation in question. The assumption that forward-looking potential customers expect a drop in prices and a rise in quality applies mostly to high-technology durables, whereas we attempt to explain the adoption of new services, which encompasses less certainty regarding the price / quality path. Furthermore, in the case of durables, customers typically pay once on the purchase occasion. Service prices, on the other hand, are periodic and in many cases changes in service price apply to existing service customers, who start paying the new periodic price when the price changes. This characteristic minimizes the value of forward-looking by potential customers. Finally, we consider service

adopters to be active market players, in the sense that they may switch between alternatives or even quit the service at any time. Customers can switch to a higher quality / lower price alternative whenever such an alternative appears; yet in the meantime obtain utility from using the service. Under such conditions, there is nothing to be gained by forward looking.

3.2 Choice models

In the choice literature, we are typically presented with models that predict choice among available alternatives as a function of alternative and decision-maker characteristics. These models usually include an option of choosing an outside alternative, or in other words "choosing not to choose". In the context of diffusion analysis, this outside alternative can be regarded as a decision not to adopt.

The framework most relevant to our form of innovation adoption and brand choice modeling, within the choice literature, is that proposed by Chintagunta and Prasad (1998). In their paper, the authors employed a dynamic McFadden model, initially proposed by Heckman and Singer (1985), to analyze inter-purchase timing and brand choice of fast-moving consumer goods. In the Chintagunta and Prasad framework, the dynamic McFadden model specified the instantaneous probability for category purchase in a given purchase occasion and the choice of a specific brand on that occasion. This probability is conditional on no purchase being made until that time (i.e., since the last purchase occasion in which a category purchase has occurred). Accordingly, the first-stage hazard function is formulated to model category purchase at a specific purchase occasion, and conditional on category purchase, a probability of brand choice is calculated.

There is, however, a significant difference between our framework and the one suggested by Chintagunta and Prasad (1998) that pertains to the nature of the investigated time intervals. Whereas inter-purchase time models focus on time intervals between consecutive purchases of fast-moving consumer goods, we concentrate on *a single specific* time interval between the time of service introduction and the time of a potential customer's initial service consideration. Naturally, this distinction evolves from the distinct nature of the products analyzed and the goals of the models.

Another recent and relevant model proposed by Krishnan et al. (2009; hereafter, KSV), suggests a two-stage purchase process that includes a time dependent category adoption process, followed by a brand choice process conditional on category adoption. This model is used to explain SUV purchase decisions of households, and is closer conceptually to our model than that proposed by Chintagunta and Prasad (1998), as it deals, in the first stage, with the process of category adoption. Conversely, in addition to the different product categories for which both models were developed (i.e., services vs. durable goods), there are two important distinctions between our proposed model and that of KSV. First, we also include a 'no-choice' option in the choice stage of our model, while KSV model accounts only for non adoption in the first-stage (i.e., no category purchase). Second, we use the first stage to model potential customers' (unobserved) service consideration decision, which is ignored by the model proposed by KSV.

Finally, a related stream of research from the choice literature are the *consideration set choice models* (e.g. Chiang et al. 1999; Gilbride and Allenby

2004; Mehta et al. 2003). In these choice models, as in our model, a decision stage that is not explicitly observed in the data is also inferred using the model. There are major differences between consideration set choice models and our model in the process underlying the unobserved decision stage, and, consequently, in the interpretation of the term *consideration* in the different models. In the unobserved process in consideration set choice models consumers are assumed to narrow down the available set of brands to a sub set for consideration, and then evaluate *only* the alternatives included in this unobserved set. In the unobserved process we model, potential customers start considering joining the service and as a result evaluate *all* service alternatives. This process may culminate in a choice of one of the alternatives, which is observed, or a decision to defer choice, i.e. no-choice. Having gone through this evaluation stage (and choosing not to choose) is what separates the no-choice but considering (NC) potential customers from the no-choice not-considering (NS) potential customers.

4 Model specification

In line with the "Dynamic McFadden" specification (Heckman and Singer 1985), we model the two-stage process of service consideration and brand choice using the combination of a hazard model and multinomial logit.

We start with the specification of λ_{it}, the transition rate from no consideration to consideration (see Fig. 1). This dynamic stage is modeled as a hazard rate. The hazard rate is a function of two elements. The first element is the length of time that has passed since the introduction of the service. This element represents the time dynamics in the diffusion process. The second element is a set of covariates that describe the market or the potential customer, and are typical to the analysis of diffusion. The hazard function covariates can, for example, be the current number of service adopters, or the periodic level of advertising. We can choose among three main alternative models for the specification of the hazard function: the Proportional Hazard Model (PHM), the Additive Risk Model (ARM), and the Accelerated Failure Time Model (AFTM—for a detailed review, see Seetharaman 2004a). In the empirical implementation of the model, the PHM specification was found to dominate the other two specifications according to the Deviance Information Criteria (DIC) and the Log Marginal Likelihood (LML) measures of fit. We therefore proceed to discuss the specification of this model.

The PHM is the most commonly used in the hazard framework (Seetharaman and Chintagunta 2003; Gonul and Srinivasan 1993; Helsen and Schmittlein 1993; Gupta 1991; Jain and Vilcassim 1991). According to the PHM, the hazard function is decomposed into two multiplicative components (we drop the customer index for convenience):

$$h_t = h_{0t} \cdot \psi(X_t). \tag{1}$$

The first component, h_{0t}, defines the baseline hazard function. This function reflects the longitudinal patterns in the duration time's dynamics. The second component is a function of $X_t = (x_{1t}, \ldots, x_{Lt})$, which is the vector of L customer and / or market covariates that affect the hazard rate. Thus, $\psi(X_t)$ adjusts h_{0t} up or down

proportionally to reflect the effect of the covariates. In most applications, $\psi(X_t)$ is formulated as an exponential function: $\psi(X_t) = \exp(\delta'X_t)$, where δ is a vector of parameters associated with the corresponding covariates affecting the proportional adjustment of the baseline hazard function. This formulation ensures the non-negativity of $\psi(X_t)$ and thereby guarantees a non-negative hazard function.

The discrete time hazard rate at time period t, λ_t, is:

$$\lambda_t = 1 - \frac{S_t}{S_{t-1}} = 1 - \frac{e^{-\sum_{u=1}^{t} \psi(X_u) \int_{u-1}^{u} h_{0w} dw}}{e^{-\sum_{u=1}^{t-1} \psi(X_u) \int_{u-1}^{u} h_{0w} dw}} = 1 - e^{-\psi(X_t) \int_{t-1}^{t} h_{0u} du}. \tag{2}$$

For the baseline hazard function, we utilize the expo-power formula as proposed by Saha and Hilton (1997) and employed by Seetharaman and Chintagunta (2003):

$$h_{0t} = \gamma \alpha t^{\alpha-1} e^{\theta t^{\alpha}}, \tag{3}$$

where γ, $\alpha > 0$.

This representation is flexible and can take many forms, including monotonically increasing, monotonically decreasing, U-shaped, or inverted U-shaped. It is also easier to implement, since it does not require numerical integration, as is the case in other specifications such as the Box-Cox formulation (Chintagunta and Prasad 1998; Jain and Vilcassim 1991).

In the choice stage, P_{jt}, the probability that a customer chooses alternative j ($j=0$, 1, …, J) at time t, conditioned on having gone through consideration, is specified as a multinomial logit model:[2]

$$P_{jt} = \frac{e^{V_{jt}}}{\sum_{j=0}^{J} e^{V_{jt}}} \tag{4}$$

V_{jt} is the deterministic part of the utility obtained from choosing alternative j at time t. It is specified to be a function of two main elements, as follows:

$$V_{jt} = \beta'Y_t + \rho_s EC_{t-1} \cdot C_{jt-1} \tag{5}$$

Where Y_t represents a set of K covariates that can characterize the alternative, the customer, or the combination of alternative-customer.

The second component in the utility function represents possible *state dependence* affecting customers' choice (Seetharaman and Chintagunta 1999; Seetharaman 2004b). This element stands for a "loyalty", "stickiness", or "status quo bias" effect (Rivot and Baron 1992; Samuelson and Zeckhauser 1988) that hypothetically raises the probability that the customer remains in the current state during the next time period. We specify this effect to influence only choice probabilities of customers who adopted the service for at least one time period by t (i.e., current

---

[2] P_{0t} is the conditional probability of opting not to choose any service alternative ($j=0$) at t $\left(\sum_{j=0}^{J} P_{jt} = 1, P_{jt} \geq 0, \forall j \right)$, given that the customer has exited the *NS* state.

adopters or quitting customers). Accordingly, we define the indicator variable EC_t as:

$$EC_t = \begin{cases} 0 & \text{if no service alternative was chosen in any time period up to and including } t \\ 1 & \text{otherwise} \end{cases}.$$

(6)

The state dependence is integrated into the model using a brand choice indicator variable:

$$C_{jt} = \begin{cases} 1 & \text{if brand } j \text{ is chosen at time } t \\ 0 & \text{otherwise} \end{cases}, \forall j = 0, 1, \ldots J.$$

(7)

Moreover, we further decompose the state effect into two differentiated effects: the "stickiness" to a service alternative (relevant to all current users of the service), and the "stickiness" to the outside option (relevant only to customers who quit the service).[3] The two state dependence effects are represented by the parameters ρ_s, $s=0,1$, where: $s = \begin{cases} 0 & \text{if } j = 0 \\ 1 & \text{otherwise} \end{cases}$.

To explain the dynamic transition of potential customers between states, at least some covariates in the two covariate sets (X of the hazard function, and Y of the utility function) must vary over time. Furthermore, we do not restrict the parameters of the two model stages to be uncorrelated, in that we allow for a correlation between the propensity to consider the service and that of choosing one of the service alternatives.

The likelihood of a customer's data in periods $t=1, \ldots, T$ is:

$$L(\overline{C}|\theta)$$
$$= \prod_{t=1}^{T} \left\{ (1 - CC_t)[Pr(NS_t|EC_{t-1}, CC_{t-1}) + Pr(NC_t|EC_{t-1}, CC_{t-1})] + CC_t \prod_{j=1}^{J} Pr(C_{jt}|EC_{t-1}, CC_{t-1}, C_{jt-1})^{C_{jt}} \right\}$$

where CC_t is a general *Choice* state indicator that signifies whether the customer uses some service alternative at time t $\left(CC_t = \sum_{j=1}^{J} C_{jt}, \text{ and } CC_0 = 0 \right)$:

$$CC_t = \begin{cases} 0 & \text{if no alternative is chosen at time } t \\ 1 & \text{otherwise} \end{cases}$$

(8)

In addition, $\Pr(NS_t)$, $\Pr(NC_t)$ and $\Pr(C_{jt})$ are the probabilities that customers reside in each of the three states—No-Service No-Consideration, No-Choice, and Choice (for a specific service alternative j), respectively—at time t. $\overline{C}$ is a vector of the choices made during $t=1$ through T, and θ is a vector of all model parameters, including the parameters of the hazard model and the multinomial logit model. The terms for the probabilities of customers residing in each of the three states, given EC_{t-1}, C_{jt-1}, and CC_{t-1}, can be obtained from the first author upon request.

---

[3] Accounting for state dependence in the no-choice option is another contribution of our model as the concept of 'stickiness' for the outside good has generally been ignored by state dependence literature.

5 Prediction

Among the more important products of any diffusion model is the prediction of future adoption. In the context of our model, once individual parameters are estimated, we want to calculate predictions of future states for each individual in the data set. To this end, we develop such future probabilities given the individuals' past histories. In other words, knowing the choice history for periods 1 through T (but not for any time after T), we develop probabilities for some future time period $T + t$. Due to the non-linearity of our likelihood function, the predictions must account for parameter uncertainty. We thus simulate probabilistic predictions for the future states, using estimated individual-level draws from the parameters' posterior distribution.

The necessary knowledge of past choices for prediction is expressed by the indicator variable EC_T. A potential customer who has never adopted ($EC_T=0$) may, at time T, be in one of the two states: No-Service No-Consideration, or No-Choice. This potential customer has a positive probability of being in all three model states at $T + t$. However, for a potential customer that has adopted the service for at least one time period by T ($EC_T=1$), the only two possible future states are Choice ($C_{jT+t}=1$ for $j>0$), or No-Choice ($C_{0T+t}=1$).

In either case, a potential customer who exits the No-Service No-Consideration state may move back and forth between the Choice and No-Choice states by re/joining and / or re/quitting the service. Thus, prediction probabilities must integrate the probabilities for every possible route between these two states, taking into account the first time period within the investigated time frame, wherein the customer exits the No-Service No-Consideration state.[4] The formulation of these probabilities and the prediction probabilities depending on EC_T can be obtained from the first author upon request.

6 Identification

The data do not distinguish non-adopters in the No Service No Consideration state from non-adopters in the No Choice state. This lack of distinction raises a potential problem of parameters identification in the two stages of our model. This potential problem is handled in two ways. First, the variables in both probability functions, (i.e., the hazard and utility functions) enter the likelihood function in rather diverse and nonlinear ways. As a result their respective parameters can be identified in the estimation procedure (Andrews and Srinivasan 1995).

Second, in the empirical application we present, different covariates enter each probability function, thus utilizing the *exclusion restrictions* approach to distinctly identify the hazard and choice probabilities. In the first (consideration) stage, covariates explain the shift to a considering potential customer and thus are similar to covariates used in the diffusion literature. Diffusion models typically employ

---

[4] When $EC_T=1$ (the simpler case), the prediction probabilities must account for every possible route between the two states, Choice and No-Choice, for the entire time frame (t time periods). When $EC_T=0$, the probability must also take into account the possibility of exiting the No-Service No-Consideration state in every future time period starting from $T+1$, and then the possible routes for the customer between Choice and No-Choice from that time period until $T+t$.

covariates that reflect the number of customers who have already adopted the innovation (internal influence) and covariates relating to external influences on diffusion. We follow this tradition in our empirical application. In the second (choice) stage, the covariates explain a considering potential customer's choice among service alternatives, including the NC option. These covariates therefore mirror covariates used in the choice literature to describe alterative/customer characteristics, or both.[5]

7 Empirical implementation

7.1 Data

The model is estimated using data on the adoption of new service plans introduced by a commercial bank to its active customers. The new service plans provided an alternative to an "old" service system used exclusively in the banking industry of the analyzed market at the time of introduction. The "old" system uses a tariff calculator for hundreds of possible transaction types. The cost for a specific transaction can range from a few cents to as much as $7. In addition to paying according to the type and volume of their activities, in the "old" system, customers also pay a fixed amount per transaction (i.e., for each row in their account balance), and a fixed monthly amount. The fixed monthly amount was often used by the bank as a discount component. Once the new service plans were introduced, each of the bank's customers had the option of continuing to use the "old" service system, or choosing among the new plans. Continuing with the "old" system is the default option, and requires no active choice on the part of customers. Therefore, in our model, this is the No-Choice alternative ($j=0$). After adopting one of the new service plans, customers can switch, or go back to the "old" system (quit in terms of our model) at any later time period (month).[6]

Customers using a plan were no longer charged according to the type and number of their transactions, but rather had periodic limits to the number of transactions they could perform free of charge in both *direct channels* and *live channels* that involve interaction with a clerk in a bank's branch or a call center. Every extra transaction above the predetermined limit entails a "penalty" payment above the basic plan fee. This additional payment is fixed and depends only on the channel used for the extra transaction.

Customers' utility from the various plans is derived not only from the number of allowed transactions made in each channel type and the periodic plan cost, but also from the clarity gained as to how much you pay and for what, as opposed to the "old" service system. Customer surveys conducted by the bank prior to the introduction of the new service plans showed that there was a lot of frustration among customers regarding the complexity and lack of transparency associated with the existing system. Moreover, customers are expected to have differing preferences

---

[5] We provide more information about the variables in the model variable section.
[6] Joining the service, switching between service alternatives and going back to the "old" system, could all be done by calling the bank branches or call center. There was no need to come in person and sign documents.

for plans partly because they perform different types and quantities of activities, and partly because *customers have differing attitudes to limits*. Neither the customers nor the bank know exactly the volume and type of future activity. There is an element of uncertainty for both parties, which is enhanced by the fact that activity may vary from period to period. Customers have *varied attitudes toward this uncertainty*. Random utility models are thus highly suitable in this context as they are derived from utility maximization premises.

The new service program was considered quite innovative for the analyzed banking industry at the time of introduction. The launch of the service was supported by a highly visible advertising campaign and received broad coverage in the national press.

We used panel data on individual adoption choices of a sample of 10,000 customers out of a list of 1,000,000 potential customers for the service. Specifically, the population of potential customers includes customers that had managed active accounts for at least six months at the time of the new service introduction. Furthermore, the information pertained only to customers defined by the bank as potential customers for the new service, and therefore does not include very young or inactive customers. The data was collected over 26 months (six months before service introduction to 20 months after introduction), and refers to every specific alternative of the new service (a total of 12 plans over the 20-month period). Figure 2 shows the new service adoption pattern during the investigated time frame.

The top line represents the total number of service adopters in each time period. The rest of the lines represent the total number of adopters for each specific plan alternative.[7] In two periods during the analyzed time frame (Periods 10 and 19), new plans were added to the existing set of plans. With the exception of one plan that was removed from the choice set in Period 10, after plans had been offered to customers, they remained available throughout the investigated time frame. Customers who had chosen the removed plan before Period 10 could still use it after its removal from the set.

7.2 Model variables

The two main variable groups in this model are the set of covariates used in the hazard model, X_t, associated with service consideration; and the logit model covariates set, Y_t, associated with brand choice. In the X_t covariate set, we have six variables:

- *Number of service adopters in the previous period* (see Fig. 2)—We expect this variable to have a positive effect on service consideration, as it represents the magnitude of possible word-of-mouth (WOM) influence of the new service (Mahajan et al. 1990).
- *Sum of marketing calls made from the bank's branches in the previous period*—This variable represents the priority level in the branches for marketing the new service compared to other marketed products.

---

[7] The graphs are not necessarily monotonically increasing, as customers can quit the service or switch to a different alternative in every time period.
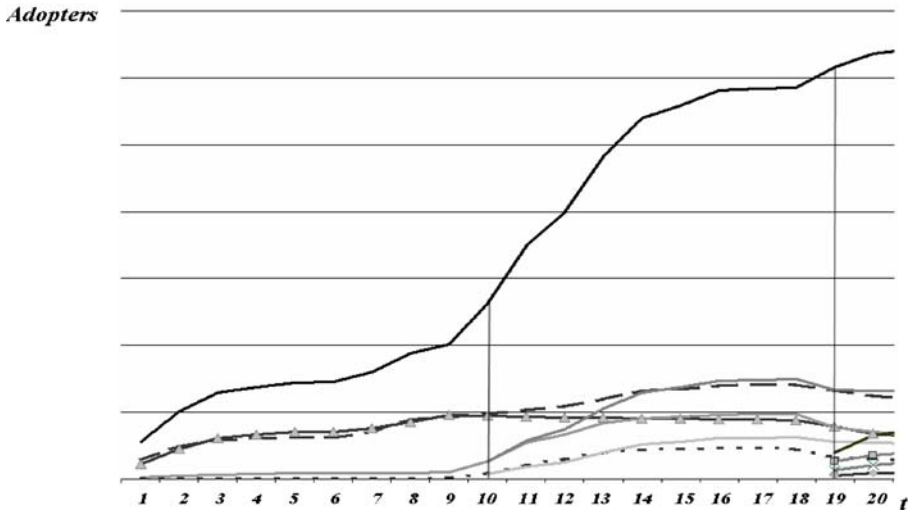
Adopters

Fig. 2 New service adoption pattern

- **Sum of marketing calls made from the bank's Call Center in the previous period**—This variable represents the priority level in the call center for marketing the new service compared to other marketed products.
- **Number of marketing calls directed at the individual customer in the previous period from the branch**—This variable represents the direct marketing effort on the part of the branch at which the customer handles the account to promote the customer's consideration of the service. The calls are expected to increase propensity to consider the new service.
- **Number of marketing calls directed at the individual customer in the previous period from the Call-Center**—although the marginal costs for these calls are lower, it is very likely that, unlike a call from the bank's branch, the customer receiving the marketing call from the call center is not familiar with the bank representative making the call. The calls from the call center are also expected to increase propensity to consider the new service.
- **Call Target indication**—An individual-level zero / one indication of whether or not the customer was on the *targeted marketing calling list*. This variable is based on the algorithm used by the bank to create the target group for marketing calls.[8]

Since the bank targeted its marketing efforts at specific customers, estimating the model without information on the targeting mechanism could potentially raise an endogeneity problem. Endogeneity may arise from the fact that the independent variables are set strategically and therefore are "endogenous" to the investigated system. We try to avoid this potential problem by integrating as much information as possible into the model, and out of the "error term" (see Manchanda et al. 2004; and Liu et al. 2007 for an elaborated discussion of the problem).

---

[8] The target indication variable is not time variant at the customer level and thus its corresponding coefficient is estimated as fixed over the whole population

Figure 3 demonstrates the aggregate number of the two types of marketing calls during the 20 months after the introduction of the innovation, and the sum of calls made in each month.[9]

From Fig. 3, we can see that the emphasis placed on the direct marketing attempts to market the service can differ greatly over months. Some months are characterized by a large volume of marketing activity promoting the adoption of the service, while in others the emphasis is probably on other products such as loans or investments options.

The bank placed these calls to introduce the new service and thereby increase service adoption; it did not attempt to trigger the adoption of any specific plan. Therefore, these variables are expected to affect the consideration of the new service, yet not the choice of a specific plan.

In the multinomial logit model, the covariates set Y_t includes seven variables:

- ▪ **New service plan intercept**—the corresponding parameter for a dummy variable, y_{0jt}, accepting the value of one for any of the new plans, and zero for the 'no-choice' option, as follows:

$$y_{0jt} = \begin{cases} 1 \text{ if } j \geq 1 \\ 0 \text{ otherwise} \end{cases}$$

This parameter represents the additional utility gained from choosing any new plan, over and above the effects of all other variables included in Y_t.

- ▪ **Calculated plan cost**—calculated for each customer for every new plan for every period according to the plan fee and the expected additional "penalty", based on the customer's activity in the previous period. The cost for the No-Choice alternative was calculated according to the number and type of transactions made by the customer in each time period. This integrates into the choice process the additional cost to the customer from any behavioral changes needed to adjust to the conditions of the plan. This variable can vary over time on the customer level and is expected to have an overall negative effect on plan choice.
- ▪ **Number of plan adopters at the previous period**—This variable is supposed to account for a possible WOM effect at the plan level. If indeed there is such an effect, it is expected to have a positive effect on choice probabilities.
- ▪ **Indication for no limit on number of transactions made through live channels**—two of the 12 plans do not restrict the number of transactions made through a clerk free of additional charge. Having no limit to the number of live transactions is expected to have a positive effect on the plan choice probabilities.
- ▪ **Indication for limited direct-channel transactions**—two of the 12 plans restrict the number of transactions made through direct channels (Web or interactive voice response) free of charge. We expect this plan attribute to have a negative effect on choice probabilities.

---

[9] For proprietary reasons we cannot identify which of the two sources is the branch and which is the Call Center.
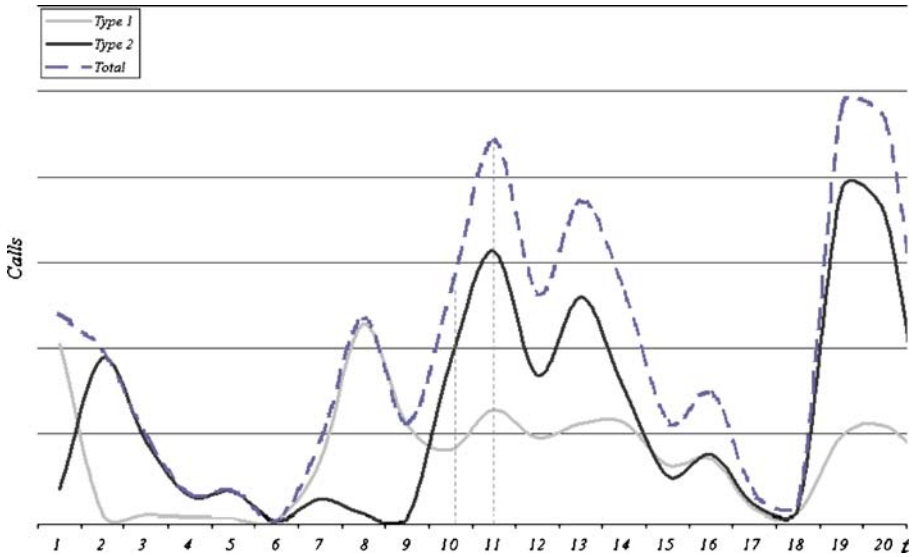
Fig. 3 Aggregate number of marketing calls

- **Other plan benefits**—an indication for additional benefits for plan users (relevant to two of the plans). These benefits are expected to increase choice probabilities.[10]
- **State dependence for plan users**—An indication for each of the new plans of whether the customer used that plan in the previous time period.
- **State dependence for service quitters**—an indication of whether the customer was in the NC state at the previous time period, after using one of the new plans for at least one time period by that time.

7.3 Estimation

For the estimation, we used the Hierarchical Bayes Markov Chain Monte Carlo (HB MCMC) algorithm that allows for a combination of random and fixed parameters (Train 2003). Allowing for fixed parameters necessitates additional layers in the Gibbs sampling of the Bayesian procedure, as will be illustrated shortly.

Let ϕ and $\{\theta_i\}$ denote the fixed parameter (coefficient of the call target indication in our model) and the population vectors of random parameters, respectively, where the index i refers to a specific customer. For the fixed parameter, we assigned a diffuse prior. For the random parameters, we assume a multivariate normal distribution with diffuse priors for the population parameters. Specifically, we use a normal prior distribution with high variance for the population means, and a diffuse inverted Wishart prior distribution for the population variance (IW (K, I), where K is the number of random parameters, and I is the identity matrix). The two

---

[10] The three plan pairs for which the last three indication variables (indication for no-limit on number of transactions made through a live channel; indication for limit on number of transactions made through direct channel; and other plan benefits) equal one are different.

parameters, α and γ, of the baseline hazard function are restricted to be positive, and are therefore exponentiated to get a lognormal distribution. The draws from the conditional posteriors for the Gibbs sampling are as follows:

First, we draw from $f\left(\{\theta_i\}|\overline{\theta}, \Sigma_\theta, \phi\right) \propto \prod_{i,t} L\left(\overline{C}|\{\theta_i\}, \phi\right) \times \Pi_1\left(\{\theta_i\}|\overline{\theta}, \Sigma_\theta\right)$ using the Metropolis-Hastings algorithm. The first element on the right-hand side is the model likelihood, and the second is the normal density. Second, θ and Σ_θ are drawn consecutively from $f\left(\overline{\theta}|\{\theta_i\}, \Sigma_\theta\right) \sim N\left(\frac{\sum_i \theta_i}{N}, \frac{\Sigma_\theta}{N}\right)$, where N is the number of customers, and from $f\left(\Sigma_\theta|\theta_i, \overline{\theta}\right) \sim IW\left(K+N, \frac{(KI+\overline{S})}{K+N}\right)$, where $\overline{S} = \sum_i \left(\theta_i - \overline{\theta}\right)$ $\left(\theta_i - \overline{\theta}\right)'$. Third, using an essentially flat prior on ϕ, we obtain draws from $f(\phi|\{\theta_i\}) \propto \prod_{i,t} L\left(\overline{C}|\{\theta_i\}, \phi\right)$ by employing the Metropolis-Hastings algorithm. The procedure consisted of 150,000 iterations, discarding the first 130,000 (burn-in iterations). We then used every tenth iteration to sample from the posterior distribution to get 2,000 draws. Overall, we have 17 parameters in the model. These include 16 random coefficients and one fixed coefficient.

7.4 Parameter estimates

Table 1 presents the estimation results for the full model. We first discuss the estimation results for the parameters in $\psi(X_t)$, the right-hand side component of the Proportional Hazard Model (PHM). These parameters are associated with the transition of potential customers from non-consideration to consideration. As expected, the number of service adopters in the previous time period affects positively the tendency of non-considering potential customers to consider the service (δ_1=17.0). Similarly, an additional marketing call to a customer, from either of the two possible channels, positively affects consideration (δ_4=47.4, δ_5=38.5). This information is of high relevance to the company as it implies to the effectiveness of the two sources as a marketing tool. There is a difference in the marginal cost of a call from the bank's branch and that from a call center representative. The assumption underlying the use of the more expensive source, the branch, is that these calls are likely to be from a person familiar to the customer and therefore should be more effective. These estimated model results can shed some more light on the overall profitability of calls from each source allowing for a better budget allocation in future direct marketing campaigns. In addition, the estimated effect of being included in the bank's target list for marketing calls is negative (δ_6=−14.1). This result is not surprising given the fact that the bank excluded from the list, among other criteria, customers who would pay less by choosing a new plan, even without changing the way they interact with the bank (i.e., without changing the number and type of transactions or the channels used to perform them). These excluded customers are likely to have the highest probability of considering the service. Finally, we found that the population of potential adopters is split with respect to the sign of total number of marketing calls effect from either channel.

We now turn to discuss the estimation results for the logit model. The intercept, β_1, indicates the tendency (utility) to adopt any new plan, over and above the effects

Table 1 Estimation results—full model

Variable	Parameter	Estimate of Population Mean Coefficient (posterior std.)[b]	Estimated Std Deviation of coefficient in population	Posterior probability [a]
Number of service adopter in previous month (1:100,000)	δ_1	17.0 (0.3)	10.5	95%
Total marketing calls in previous month (Type 1) (1:10,000)	δ_2	0.3 (0.2)	3.2	54%
Total marketing calls in previous month (Type 2) (1:10,000)	δ_3	0.4 (0.3)	4.3	54%
Number of marketing calls to the customer in previous month (Type 1)	δ_4	47.4 (1.5)	32.8	93%
Number of marketing calls to the customer in previous month (Type 2)	δ_5	38.5 (1.5)	36.4	85%
Call target indication	δ_6	−14.1 (0.6)	–	–
Intercept	β_1	−4.9 (0.5)	20.8	41%
Calculated plan cost	β_2	−0.5 (0.02)	0.5	16%
Number of plan adopters in previous month (1:10)	β_3	0.3 (0.1)	1.7	58%
Unlimited live-channel transactions	β_4	6.1 (0.5)	8.8	76%
Limited direct-channel transactions	β_5	−11.5 (0.6)	7.9	7%
Other plan benefits	β_6	−0.4 (0.9)	25.9	49%
State dependence for plan users	ρ_1	9.9 (0.4)	4.5	99%
State dependence for service quitters	ρ_0	10.3 (0.4)	8.6	88%

[a] The posterior probability that the coefficient is greater than zero

[b] The numbers in brackets represent the posterior standard deviation for the population parameter mean

of other variables. Although on average it is negative (−4.9), indicating basic resistance to change, 41% of the posterior distribution exhibits a positive tendency to adopt a new plan (conditioned on consideration). The calculated plan cost, representing the expected sum to be paid for the plan according to the customer's activity characteristics, has a negative effect on plan choice (β_2=−0.5). The effect of using a specific plan in the previous time period on the probability of continuing to use it is positive (ρ_1=9.9) and is equivalent to almost 5 USD (ρ_1/β_2), compared to a basic plan fee range of 5 to 12 USD. Similarly, being in the NC state in the previous time period, for customers who quit the service, has a positive effect on the probability of remaining in that state (ρ_0=10.3). Although both state variables have similar estimated mean effects, the standard deviation of the state effect parameter for service quitters is almost twice that of the state effect for plan users. The greater variance indicates a higher heterogeneity in the tendency to avoid rejoining the service for customers who quit the service, compared with the tendency of service users to stay with their current plan.

Having a limit on the number of transactions in both live and direct channels has a negative effect on choice probabilities for that plan (β_4=6.1, β_5=−11.5). For most customers it is more important not to have a limit on the number of direct channel

transactions, than it is to have no limit on live channel transactions. In addition, we find that customers who consider joining the service are split in the way the number of plan adopters and the additional plan benefits affect their probability of choosing the plan (β_3=0.3, Std Deviation=1.7, β_6=−0.4, Std Deviation=25.9).

We also investigated the correlations between hazard parameters and the intercept in the logit model, β_1 (Chib et al. 2004). The intercept represents the utility from adopting a new plan beyond that gained from all other plan or customer characteristics. Its correlation with hazard parameters may reveal a possible linkage between the tendency to consider the service and the tendency to choose one of the new service plans. We find that the logit intercept has a relatively high correlation with δ_1, the effect of previous period adopters on consideration rate. The correlation between β_1 and δ_1 is (−0.7). This means that customers who are less influenced by WOM have a higher tendency to choose a plan once they consider the service. Moreover, in a factor analysis, conducted on the individual parameter means, we get that the two parameters β_1 and δ_1 also have high loadings on the same factor (factor loadings of 0.90 and −0.78 for δ_1 and β_1, respectively). Due to space limitations, we do not present the entire correlation table or the factor analysis pattern, but rather generally report the factor analysis results.[11] The procedure identified 3 factors that account for 92.36% of the system variance. Customers with high factor scores for the first factor can be characterized as more '*careful*' customers when it comes to adopting a new service plan. They are more sensitive to the overall 'atmosphere' surrounding the new plans—i.e., the total number of plan adopters, and the volume of Type 1 marketing calls conducted (δ_1, and δ_2). They are less inclined to adopt a plan once they start considering the new service plans (β_1), and are more sensitive to the activity characteristics of the plans (β_4, β_5, and β_6). They are also more price sensitive (β_2). Customers with high factor scores for the second factor are more positively affected by the marketing calls they receive (δ_4 and δ_5), and are more sensitive to the volume of Type 2 marketing calls (δ_3). These customers are also more inclined to keep away from the new plans after trying them and quitting the new service (ρ_0). Customers with high factor scores for the third factor have a higher '*plan loyalty*' in the sense that they have a higher tendency to stick with the plan they adopt (ρ_1), and are also more sensitive to the number of adopters for a plan in the choice process (β_3).

7.5 Additional insights

In addition to offering insights into the covariates that affect the flow between adoption states, the model separates the non-adopter population into two segments: No-Service No-Consideration (NS) customers, and No-Choice after-consideration (NC) customers. This partition suggests important implications about the diffusion process on both the aggregate (evaluating the two segment sizes) and the individual (customer-level indicators of pre-adoption stages) levels. After estimating the parameters, we calculated the estimated probabilities of being in NS and NC for

---

[11] The full correlation table and the results of the factor analysis are available upon request from the first author.

each non-adopter, and then assigned a non-adopter state by choosing the state with the highest estimated probability. Figure 4 shows the estimated percentage of non-service users in both states.

The percentage of NC customers increased from less than 1% to more than 19% over the investigated time frame, indicating that more non-adopters started considering the service, yet did not find a plan that suits them. The jump in the percentage of NC potential customers in Periods 10, 11, and 12 can be explained by the high number of marketing calls directed at customers in Periods 10 and 11, and the increase in adoption rates starting from Period 9. Nonetheless, the vast majority of potential customers remained in NS throughout the investigated time frame. This indicates that most non-adopters did not even consider joining the service. Therefore, the more prominent action needed in order to encourage adoption among non-adopters should be emphasizing those factors that have a significant effect on the hazard function. These factors include directing marketing calls to non-considering potential customers. In addition, once more customers adopt the service as a result of these marketing calls, the increase in number of adopters will in turn also have a positive effect on consideration rates. For non-adopters in NC, the insights for the model demonstrate that the recommended approach is to manipulate those factors that trigger choice (e.g., plan price and assigning a limit on the number of transactions in either of the channels). We further see from the estimation results that the additional benefits received in two of the plans do not play an important role in plan choice. One course of action would be to consider other benefits to encourage plan adoption. Such plan benefits were introduced in Period 19 with the introduction of new plans.

Since probabilistic indications as to non-adopters' states are also given at the individual customer level, the model provides an important marketing tool that companies can use to target marketing strategies and tactics at specific customers and thus advance the diffusion process.
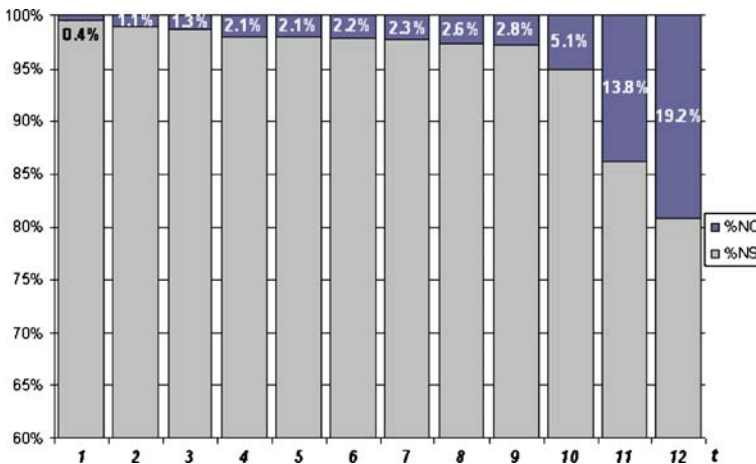

Fig. 4 Partition of non-service users into NS and NC States

7.6 Evidence from simulated data

In order to evaluate our model's ability to distinctly identify the two pre-adoption states we conducted several simulated experiments. In each experiment we have constructed a data set based on the actual data of our samples' customers. Individual parameters were randomly drawn from a population distribution and were assigned to each customer in our sample. Then, using the variables of the customer in our actual data, we calculated the three state probabilities (i.e., NS, NC and CC) for every customer. Accordingly, by assigning each customer to the state with the highest probability (and likewise to the plan with the highest probability for customers in CC), we have constructed a simulated service plan adoption variable that indicates for every time period whether the customer adopted a new plan and if so what is the specific plan chosen based on our model. In order to evaluate the model's ability to identify the latent non-adopter groups under different adoption patterns, each simulation experiment differed in the population parameter distributions (i.e., population mean and variance). We then used the simulated data in a Bayesian estimation procedure similar to that used in the empirical application, to retrieve the individual parameters. These in turn were employed to compute the individual model states probabilities.

Once we have assigned a 'post estimation' state to each non adopter (NS or NC), based on the state with the highest probability, we compared the original simulated states to those obtained after the estimation. Across all 5 experiments we have achieved a good 'hit' rate for the predicted unobserved states. On average, over the 5 experiments and over the 12 estimation periods, 94.5% of all customers-periods with simulated NS state were correctly assigned to the NS state based on the estimation results. Similarly, 86.2% of the customers originally belonging to the NC group were assigned correctly to the NC state once the estimated individual parameters were obtained.[12]

7.7 Model comparison

We compare our full model to three alternative models estimated using the same data set. This comparison has three main goals. First, to compare between the quantity and quality of information and insights derived from each model. The second goal is to compare models' fit; and finally, through the additional estimations, we can compare models' predictive ability.

The chosen alternative models represent commonly used models to analyze either adoption or choice processes. The first is an individual-level (mixed) PHM with an expo-power specification for the baseline hazard function. The PHM integrates the same X_t covariate set as the full model, and represents an individual-level binary diffusion model. Second, we estimate two individual-level choice models using the mixed multinomial logit. These models represent a static choice process for plan adoption and choice. The first choice model (ML1) includes the same covariate set used in the choice stage of the full model. The second choice model (ML2) includes,

---

[12] Due to space limitations, we do not present the entire simulation state assignment tables. These tables are available upon request from the first author.

in addition to the covariates in ML1, the covariates that enter the consideration stage through the hazard function of the full model. The additional covariates in ML2 enter the utility function of all new plans, yet not that of the no-choice option. In this way, we integrate into the choice system information embedded in covariates that are expected to affect customers' inclinations to consider adopting the new service.

The three individual-level models are estimated using the same HB MCMC procedure as we applied to the full model, and we again base the estimation on Periods 1–12. Table 2 reports the estimation results of the three models together with those of the full model.

We discuss first the estimation results for the parameters of $\psi(X_t)$ in the PHM. Like the estimated results of the full model, here also, we find a strong positive effect of number of adopters on service adoption ($\delta_1 = 25.8$). Similarly, there is a positive estimated effect of the number of marketing calls a customer received on adoption ($\delta_4 = 37.0$ and $\delta_5 = 20.2$), although we see higher heterogeneity in these parameters' distributions in the PHM compared to those estimated in the full model. The main difference between the estimated results of the PHM and those of the hazard parameters in the full model centers on the effect of the total number of marketing calls. Unlike our expectations, the two parameters corresponding to these covariates have a negative estimated mean ($\delta_2 = -1.3$ and $\delta_3 = -2.5$), with a substantial part of the posterior distribution of the these parameters being negative. We cannot come up with an explanation for this outcome.

We turn now to discuss the estimation results of the two alternative choice models, ML1 and ML2. In both models, we see a negative sign for the choice intercept for the vast majority of the population ($\beta_1 = -2.8$ and $\beta_1 = -33.3$ in ML1 and ML2 respectively). The intercept in the logit model represents the utility derived from choosing a new service plan, in addition to all other factors that influence the utility from a plan. In the two choice models, the negative estimated intercept indicates a basic negative tendency toward adopting the new service. The full model reveals a much more heterogeneous distribution for the intercept, with 41% of the posterior distribution being positive. This difference may hint at the pure logit model's lack of ability to capture the time-dependent element associated with adoption, an element that is taken into account in the full model. As a result, the intercept in the choice models must express both the Consideration and the Choice components in customers' tendencies to defer choice to later time periods and not adopt the service right after its introduction.

In addition, in both choice models, we unexpectedly got a negative effect of having no limit to the number of transactions made through live channels on plan choice probabilities ($\beta_4 = -3.4$ and $\beta_4 = -4.7$ in ML1 and ML2 respectively). A similar unexpected negative effect was found in both choice models for other plan benefits ($\beta_6 = -2.9$ and $\beta_6 = -9.7$ in ML1 and ML2 respectively).

The estimation results of ML2 reveal several other results that contradict both prior expectations and the results of the other estimated models. These results include a negative effect of previous period number of adopters on new service plan adoption ($\beta_{11} = -0.6$), and a negative state effect for service quitters ($\rho_0 = -19.8$). In addition, unlike the results of the full model and the ML1, we do not get a conclusive plan cost effect in ML2 ($\beta_2 = -0.1$, Std Deviation of 0.6).

Table 2 Estimation results alternative models

Variable		Full Model Estimate (std.) [b]	Posterior prob. [a]	PHM Estimate (std.) [b]	Posterior prob. [a]	ML1 Estimate (std.) [b]	Posterior prob. [a]	ML2 Estimate (std.) [b]	Posterior prob. [a]
Number of service adopter in previous month (1:100,000)	δ_1/β_{11} [c]	17.0 (0.3)	95%	25.8 (0.8)	98%	–	–	−0.6 (0.04)	14%
Total marketing calls in previous month (type 1) (1:10,000)	δ_2/β_{12} [c]	0.3 (0.2)	54%	−1.3 (0.2)	34%	–	–	0.4 (0.1)	64%
Total marketing calls in previous month (type 2) (1:10,000)	δ_3/β_{13} [c]	0.4 (0.3)	54%	−2.5 (0.1)	13%	–	–	−0.4 (0.1)	31%
Number of marketing calls to the customer in previous month (type 1)	δ_4/β_{14} [c]	47.4 (1.5)	93%	37.0 (1.4)	78%	–	–	21.7 (0.7)	90%
Number of marketing calls to the customer in previous month (type 2)	δ_5/β_{15} [c]	38.5 (1.5)	85%	20.2 (1.3)	67%	–	–	13.8 (0.8)	78%
Call target indication	δ_6/β_{16} [c]	−14.1 (0.6)	–	−6.3 (1.1)	–	–	–	−0.5 (0.4)	–
Intercept	β_1	−4.9 (0.02)	41%	–	–	−2.8 (0.1)	6%	−33.3 (1.5)	6%
Calculated package cost	β_2	−0.5 (0.2)	16%	–	–	−0.2 (0.01)	18%	−0.1 (0.01)	45%
Number of package adopters in previous month (1:10)	β_3	0.3 (0.5)	58%	–	–	−3.1 (0.3)	25%	1.2 (0.2)	69%
Unlimited 'human' channel transactions	β_4	6.1 (0.6)	76%	–	–	−3.4 (0.6)	31%	−4.7 (0.3)	31%
Limited direct channel transactions	β_5	−11.5 (0.9)	7%	–	–	−9.2 (0.7)	5%	−4.6 (0.2)	24%
Other package benefits	β_6	−0.4 (1.3)	49%	–	–	−2.9 (0.3)	33%	−9.7 (0.7)	31%
State effect for package users	ρ_1	9.9 (0.4)	99%	–	–	9.9 (0.3)	100%	39.6 (1.2)	100%
State effect for service quitters	ρ_0	10.3 (0.1)	88%	–	–	15.0 (0.8)	99%	−19.8 (1.7)	25%

[a] The posterior probability is the probability that the parameter is greater than zero.

[b] The numbers in brackets represent the posterior standard deviation for the population parameter mean

[c] In ML2 these variables enter the utility function of new plans

The estimation results of ML2 allow us to investigate the sensitivity of our model to parametric assumptions. Note that the combination of hazard and choice in the dynamic McFadden framework essentially relaxes the logit model's restriction that the "error term" for the outside alternative has the same distribution as the errors for the J brands, and that this distribution is constant over time. In our setting, the integration of the additional hazard function in effect allows a mass point at negative infinity in the logit model, as long as no brand has yet been chosen. The full model and ML2 make use of the same variables through differing parametric assumptions. Therefore, the comparison between the estimation results of the two models, and specifically the multitude of unexpected findings in ML2, provides additional support for the behavioral assumptions underlying our full model.

To conclude, the estimation results of the alternative models reveal many unexpected results, suggesting lower face validity for these models and hinting at their lack of ability to capture all processes that take place in the data.

7.8 Model fit

To compare the fit of the models, we employed two measures often used for Bayesian model selection: the Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002), and the log marginal density (LMD) (Newton and Raftery 1994). Since the PHM models only the binary adoption (yet not the specific plan choice), we also calculate two alternative measures of fit for the full and the choice models using only binary adoption data. This calculation enables comparison between the fit of the three brand-level models and that of the binary PHM. Table 3 presents the fit measures.

The full model dominates the two choice models (ML1 and ML2) by DIC and LMD in both the brand and the binary level. So, combining a hazard model and a choice model (as is done in the full model) provides better model fit than that provided by a choice model alone, even if one includes in the choice model all variables relevant to choice and consideration, as is the case in ML2. Comparing the full model with the binary PHM, we see that it is worse by LMD. According to the DIC measure for the PHM, we get an unusually low (negative) value. The odd figure results from very high correlations between several pairs of the model parameters.[13] This correlation structure causes some individuals to have low likelihood values (and accordingly very low log-likelihood values) when using the parameters' mean, while having reasonable likelihood values for each draw from the posterior distribution. This produces a large difference between the mean of the log-likelihood values for the 2,000 draws, and the log-likelihood value of the parameter means, i.e., the two components used for the DIC calculation. This result is another indication of the difficulty in estimating individual-level parameters for an adoption process using the PHM without augmenting it with a choice process.

---

[13] Due to space limitations, we do not present the whole correlation table. This table is available upon request from the first author.

Table 3 Models fit comparison: DIC and LMD

	Full Model	PHM	ML1	ML2
DIC	13,602	–	35,730	17,184
Binary DIC	2,440	−96,654	22,617	7,069
Log marginal density	−6,847	–	−15,858	−6,901
Binary log marginal density	−1,687	−1,285	−9,832	−1,820

7.9 Model predictive ability

To examine the predictive ability of the model, we use the holdout sample of eight time periods (13–20). Due to the nonlinearity of the likelihood function, we base our prediction calculations on all 2,000 draws from the posterior distribution of the model parameters for a random sub-sample of 1,000 customers. That is, we simulate future state probabilities across all holdout periods for each of the sub-sample customers, by employing the parameters of every draw, and then average out the results (for each period / customer) over the 2,000 calculations. Next, we compare these predictions with the actual choices depicted in the data. We calculate the predictions of the four estimated models, i.e., the full model, PHM, ML1, and ML2.

In Period 19, five new plans were introduced (i.e., after parameter estimation). Each of the five new plans is similar to one of the existing plans in terms of the number of allowed transactions and all other plan benefits, yet is more costly and offers the option of account-level interest offsetting, a newly introduced plan feature for which the procedure does not estimate the derived utility. We therefore calculated the predicted probability of joining a new plan only based on the set of characteristics for which we estimated the corresponding parameters. This technique is expected to generate underestimated prediction probabilities for the new plans if the new offsetting adds to customers' utility more than does the decrease in utility from the higher plan price, and will create overestimated probabilities in the opposite case.

We first discuss the overall predicted diffusion pattern according to the four models. Figure 5 presents the calculated adoption rates and the actual adoption for all holdout periods. From Fig. 5, we can see that although it does not depict the full magnitude of adoption in future time periods, the full model's predicted pattern of adoption is the closest to the actual pattern in terms of shape and adoption numbers. The PHM dramatically overestimates adoption, to as much as 324% of actual adoption in some holdout periods. This over prediction for adoption probably results from a combination of reasons. First, the PHM is based on a limited number of variables that predict adoption compared with the full model. The predictions of this model are therefore highly sensitive to changes in these variables. In addition, as opposed to the full model, adoption in the PHM is determined based on the hazard rate alone. In our dataset, three of the five variables modeled to affect the hazard rate, display a significant change starting from the 15 time period—the point in time where the PHM predictions start to overestimate adoption. First, there is a sharp increase in the number of adopters at the 14th time period (see Fig. 2), and, at the same time, a sharp drop in the total number of marketing calls from both sources
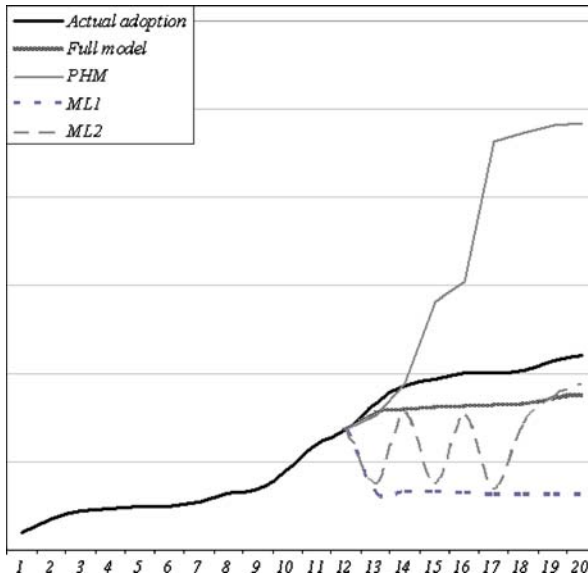
Fig. 5 Actual and predicted adoptions rates for holdout periods

(see Fig. 3). Given the estimated sign of the corresponding parameters (i.e., δ_1, δ_2, and δ_3), we get a sharp increase in the predicted hazard rates, and therefore an overestimation of adoption.

The two pure choice models (ML1 and ML2), on the other hand, tend to underestimate future adoption rates. This result is due to the significant negative intercept estimated in both models, demonstrating the inability of the choice-only logit model to capture the time-dependent element associated with service adoption.

We further see that ML2 (a choice model that incorporates the covariate set of the hazard stage in the full model) reveals an unstable adoption pattern in the first six holdout periods. The unstable predictions result from the combination of a very low intercept and a negative state effect for the NC state. This combination can cause reverse adoption predictions in subsequent prediction periods. Starting from Period 18, the predicted adoption pattern of ML2 stabilizes, probably due to a positive effect of an increase in total number of adopters and the number of marketing calls made during these time periods.

For the overall diffusion pattern prediction we also tried to aggregate the data and empirically compare the full models projections with those of other brand-level models existing in the literature. Specifically, the models suggested by Parker and Gatignon (1994) and by Krishnan et al. (2000), are relevant to a diffusion setting wherein there are multiple brands introduced sequentially. These models, being based on aggregate data, require a larger number of observation points in order to estimate the full set of model parameters. Specifically, Parker and Gatignon (1994) used 26 data points for each of the six brands studied, and KBK (2000) used 56 data points for the incumbent brands and 12 data points for the new entrant. In addition, the estimation of the Parker and Gatignon model requires information on advertising data that we do not have, yet can replace with the direct marketing calls data.

The data we have to compare the estimation results and predictions of the full model and these aggregate brand-level diffusion models, contains 12 data points for the plans that were available right from service introduction, and only three data points for the two brands that were introduced in Period 10. This quantity of data points does not allow the estimation of all brand-level model parameters. We also tried to estimate the KBK model using 13 and 14 time periods, over which there are just enough data points for the estimation, yet were not able to reach convergence for our data.

We also compared the models on accuracy of predicted adoption on the individual level. This comparison offers important database marketing implications, because it reflects the models' ability to pinpoint customers who are likely to adopt in future time periods and toward whom marketers should direct their efforts in order to start benefiting from such customers' adoption in *earlier time periods*. A good model, from this perspective, should offer the highest correct adoption predictions, and therein enable a more targeted and efficient spending of the marketing budget.

It is important to note that in our empirical example, the direct marketing efforts were based on economic criteria alone and not on measures of call success likelihoods. That is, the list of bank customers to receive a direct marketing call consisted of accounts holders that were expected to benefit from adopting a package only by changing the way they handle their accounts (i.e., by shifting more activities to direct channels). Our model, combined with the existing list construction mechanism, can provide information on *adoption probabilities* and thus offer more dimensions to prioritize calls.

For each model, we constructed a group of customers who were not using the service at Month 12, yet are predicted to use the service in Month 20 (we refer to this group as the "target group").[14] In this way, we can compare predictions eight periods ahead. We then calculated the percentage of customers in this group that actually used the service in Month 20. Table 4 presents the relative size of the group predicted to adopt and the percentage of correct adoption predictions for each model.

The full model predicts that of the non-adopters in Period 12, 6% will be using one of the service plans by Period 20. As it turned out, 78% of these actually were using the service. As can be seen from the table, this is the highest percentage of correct adoption predictions compared to all other models. The full model is followed by ML2, which, probably due to the stabilization in the prediction pattern starting in Period 18, offers relatively high percentages of correct adoption predictions (72%). The PHM predicts a very high adoption rate (95%), yet only 24% of these customers actually adopted the service. ML1, on the other hand, predicts that none of the customers in the non-adopter population will adopt in future time periods.

7.10 Prediction analysis for customers contacted by the bank

In order to measure the accuracy of the model at predicting customers' responsiveness to marketing activities we conducted a separate examination of the

---

[14] We chose the 20th time period as the basis for Table 4 since it is the last time period in our data, and therefore allows us to demonstrate the model's predictive ability for a relatively far time period. The prediction results are never the less consistent for earlier time periods as well.

Table 4 Relative size of target group and correct adoption prediction percentages

	Target group % out of non-adopter population in Period 12	Correct adoption prediction percentages
Full Model	6%	78%
PHM	95%	24%
ML1	0%	–
ML2	10%	72%

customers that have not adopted by the 12th time period and received a marketing call at some point during the prediction time frame (i.e., between the 12th and the 20th periods). According to our model, among this group of customers, 13% are predicted to adopt the service by the 20th time period. The actual adoption rate of these customers is 78%, while the actual adoption rate in the remaining 87% of customers receiving a call in this time frame is only 32%. Moreover, among customers who received a marketing call and eventually adopted, the mean probability to adopt is much higher than that for non adopters (mean adoption probability of 0.75 compared with 0.29 for eventual adopters and non adopters, respectively). This outcome is interesting as it indicates that the model also enables us to recognize among the customers receiving a call which are more likely to adopt, and accordingly to prioritize the order of calls. Note that the list of customers that received a marketing call was not constructed according to adoption probabilities, but rather according to the bank goals in offering the new service plans. The model thus adds valuable information that can help managers conduct marketing activities more efficiently.

8 Discussion and future research

The motivation for this research was raised while searching for an analytic tool that considers simultaneously both the size of the diffusion "pie" and the factors affecting its partition into various "wedges". While models are available in the marketing literature for the analysis of each element separately, the availability of models that deal with both is quite limited, particularly at the individual customer level. Typical diffusion models do not deal with choice processes that are inherent to the many adoption processes and their effects on the diffusion pattern. Choice models in the context of new products / services, on the other hand, do not integrate a time element, and thus fail to reflect an important process at the heart of new products / services adoption.

In this paper, we present an individual-level model of diffusion and choice. The two-stage nature of this model enables the representation of two consecutive sub-processes that comprise the adoption process: service consideration and brand choice.

The model adds to the scarce literature on individual and brand level diffusion models, and allows us to no longer regard all non-adopters as part of one

homogeneous pool of potential market. Instead, the estimation of the model, using adoption data of a new service, facilitates the identification of potential customers from differing stages of the diffusion process. Moreover, the model enables us to identify the factors that have a significant influence on both types of potential customers, and to assess possible associations between the effects of these factors.

When we compare the performance and products of our proposed model with those of alternative existing models from either the diffusion or the choice literature, we see that in addition to a broader analysis (i.e., the notion of pre-adoption groups and the brand-level analysis), the model also provides superior fit and prediction accuracy.

In addition, the model also relates to the CRM literature with the insights it provides on different stages in the customer lifecycle. Unlike most product categories analyzed in the diffusion literature, the focus here is on the diffusion of a new service taking into account specific behavioral characteristics that are relevant also after the initial adoption of the service. These characteristics include active switching between service alternatives and the possibility to quit the service all together. Taking these behaviors into account allows us to draw insights that are broader than the limited pre-adoption perspective. Specifically, we no longer focus merely on customer acquisition but also expand the analysis framework to retention and development (Blattberg et al. 2001). Using information provided by the model on the likelihood of adoption together with that of switching to another service alternative or of quitting the service, managers can improve their ability to optimize marketing spending (direct and indirect) over different stages in the customer lifecycle.

Our two-stage-three-states model rests on the assumption that people, who went through the evaluation process but could not find any alternative that will improve their present situation, and therefore decided not to adopt (NC state), are fundamentally different from those who did not go down the evaluation path at all (NS state). We called this path (process) "consideration". All people who ended up adopting one of the new service alternatives went through the consideration process, but not all people who considered ended up adopting the service. The consideration process requires investment of time and effort for collecting and processing information about the merits of the various alternatives offered. Before doing so, people in the NS state have some expectation about the possible gain from this process. Naturally, at this stage people have a rather vague idea about the possible gains, so the variance around their expectation is rather large. Some people find this expectation too small and/or the variance too large to be worth the investment, so they opt not to enter the consideration process. Those who perceive the odds to be favorable take the chance and invest the necessary resources to evaluate the alternatives. Among them, some do not find any alternative that is worth their while. These people may resume the evaluation process very easily with every change in the market without having to invest all the resources that go into a full fledged evaluation. They are relatively easy prey for marketing effort.

In the present paper we modified well known models to construct our individual level service diffusion and brand choice model. One could try to develop such a model from overall utility maximization principles. For example, investment in consideration (time and effort) could be viewed as means to reduce the variance of

the gain expected from an alternative. Risk attitude maybe incorporated in the utility function. This way expected utility conditioned on consideration will differ from unconditional expected utility of people in the NS state. We leave the challenge of such formal development to future research.

References

Andrews, R. L. & Srinivasan, T. C. (1995). Studying consideration effects in empirical choice models using scanner panel data. *Journal of Marketing Research, 32*, 30–41.

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science, 15*, 215–227.

Blattberg, R. C. Getz, G. & Thomas, J. S. (2001). *Customer equity: building and managing relationships as valuable assets*. Boston: Harvard Business School Press.

Bolton, R. N. (1998). Dynamic model of the duration of the customer's relationship with a continuous service provider: the role of satisfaction. *Marketing Science, 17*(1), 45–65.

Berry, L. L. (2002). Relationship marketing of services: perspectives from 1983 and 2000. *Journal of Relationship Marketing, 1*(1), 59–70.

Chatterjee, R. & Eliashberg, J. (1990). The innovation diffusion process in a heterogeneous population: a micromodeling approach. *Management Science, 36*, 1057–1080.

Chatterjee, R. Eliashberg, J. & Rao, V. R. (2000). Dynamic models incorporating competition. In V. Mahajan, E. Muller & Y. Wind (Eds.), *New product diffusion models*. New York: Kluwer Academic.

Chiang, J. Chib, S. & Narasimhan, C. (1999). Markov chain Monte Carlo and models of consideration set and parameter heterogeneity. *Journal of Econometrics, 89*(1–2), 223–248.

Chib, S. Seetharaman, P. B. & Strijnev, A. (2004). Model of brand choice with a no-purchase option calibrated to scanner-panel data. *Journal of Marketing Research, 41*(2), 184–196.

Chintagunta, P. K. & Prasad, A. R. (1998). An empirical investigation of the dynamic McFadden model of purchase timing and brand choice: implications for market structure. *Journal of Business and Economic Statistics, 16*, 2–12.

Dodson, J. & Muller, E. (1978). Models of new product diffusion through advertising and word of mouth. *Management Science, 24*, 1568–1579.

Gilbride, T. J. & Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science, 23*(3), 391–406.

Givon, M. Mahajan, V. & Muller, E. (1995). Software piracy: estimation of lost sales and their impact on software diffusion. *Journal of Marketing, 59*, 29–37.

Givon, M. Mahajan, V. & Muller, E. (1997). Assessing the relationship between the user-based market share and unit sales-based market share for pirated software brands in competitive markets. *Technological Forecasting and Social Change, 55*, 131–144.

Gonul, F. & Srinivasan, K. (1993). Consumer purchase behavior in a frequently bought product category: estimation issues and managerial insights from a hazard function model with heterogeneity. *Journal of the American Statistical Association, 88*, 1219–1227.

Gupta, S. (1991). Stochastic models of inter-purchase time with time-dependent covariate. *Journal of Marketing Research, 28*, 1–15.

Heckman, J. & Singer, B. (1985). *Longitudinal analysis of labor market data*. Cambridge, UK: Cambridge University Press.

Helsen, K. & Schmittlein, D. C. (1993). Analyzing duration times in marketing evidence for the effectiveness of hazard rate models. *Marketing Science, 11*, 395–414.

Hogan, J. E. Lemon, K. N. & Libai, B. (2003). What is the real value of a lost customer? *Journal of Service Research. 5*(3), 196–208.

Horsky, D. (1990). A diffusion model incorporating product benefits, price, income, and information. *Marketing Science,* (Fall) 342–365.

Jain, D. & Vilcassim, N. J. (1991). Investigating household purchase timing decisions: a conditional hazard function approach. *Marketing Science, 10*, 1–23.

Jain, D. Mahajan, V. & Muller, E. (1991). Innovation diffusion in the presence of supply restrictions. *Marketing Science, 10*(1), 83–90.

Kalish, S. (1985). A new product adoption model with price, advertising, and uncertainty. *Management Science, 12*, 1569–1585.

Krishnan, T. V. Bass, F. M. & Kumar, V. (2000). Impact of a late entrant on the diffusion of a new product / service. *Journal of Marketing Research, 37*, 269–278.

Krishnan, T. V., Seetharaman, P. B., & Vakratsas, D. (2009). "Size" versus "Slice" of the pie: modeling category- and brand-level SUV sales. Working paper.

Lilien, G. L. Rangaswamy, A. & Van den Bulte, C. (2000). Diffusion models: managerial applications and software. In V. Mahajan, E. Muller & Y. Wind (Eds.), *New product diffusion models*. New York: Kluwer Academic.

Liu, Q. Otter, T. & Allenby, G. M. (2007). Investigating endogeneity bias in marketing. *Marketing Science, 26*(5), 642–650.

Mahajan, V. Muller, E. & Kerin, R. A. (1984). Introduction strategy for new products with positive and negative word of mouth. *Management Science, 12*, 1389–1405.

Mahajan, V. Muller, E. & Bass, F. (1990). New product growth models in marketing: a review and directions for research. *Journal of Marketing, 54*(1), 1–26.

Manchanda, P. Rossi, P. & Chintagunta, P. (2004). Response modeling with non-random marketing mix variables. *Journal of Marketing Research, 41*(4), 467.

Mehta, N. Rajiv, S. & Srinivasan, K. (2003). Price uncertainty and consumer search: a structural model of consideration Set Formation. *Marketing Science, 22*(1), 58–84.

Muller, E., Peres, R., & Mahajan, V. (2009). Innovation diffusion and new product growth: Critical review and research directions. Working paper.

Newton, M. A. & Raftery, A. E. (1994). Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society, Series B (Methodological), 56*(1), 3–48.

Parker, P. & Gatignon, H. (1994). Specifying competitive effects in diffusion models: an empirical analysis. *International Journal of Research in Marketing, 11*, 17–39.

Reinartz, W. J. & Kumar, V. (2000). On the profitability of long-life customers in a non-contractual setting: an empirical investigation and implications for marketing. *Journal of Marketing, 64*, 17–35.

Rivot, I. & Baron, J. (1992). Status quo omission bias. *Journal of Risk and Uncertainty, 5*, 49–61.

Roberts, J. H. & Lattin, J. M. (2000). Disaggregate-level diffusion models. In V. Mahajan, E. Muller & Y. Wind (Eds.), *New product diffusion models*. New York: Kluwer Academic.

Roberts, J. H. & Urban, G. L. (1988). Modeling multiattribute utility risk and belief dynamics for new consumer durable brand choice. *Management Science, 34*(2), 167–186.

Saha, A. & Hilton, L. (1997). Expo-power: a flexible hazard function for duration data models. *Economics Letters, 54*, 227–233.

Samuelson, W. & Zeckhauser, R. (1988). Status quo bias in decision-making. *Journal of Risk and Uncertainty, 1*, 7–59.

Seetharaman, P. B. (2004a). The additive risk model for purchase timing. *Marketing Science, 23*(2), 234–242.

Seetharaman, P. B. (2004b). Modeling multiple sources of state dependence in random utility models: a distributed lag approach. *Marketing Science, 23*(2), 263–271.

Seetharaman, P. B. & Chintagunta, P. K. (1999). Investigating household state dependence effects across categories. *Journal of Marketing Research, 36*(4), 488–500.

Seetharaman, P. B. & Chintagunta, P. K. (2003). The proportional hazard model for purchase timing: a comparison of alternative specifications. *Journal of Business Econometrics and Statistics, 21*(3), 1–15.

Song, I. & Chintagunta, P. K. (2003). A micromodel of new product adoption with heterogeneous and forward-looking consumers: application to the digital camera category. *Quantitative Marketing and Economics, 1*, 4–371.

Spiegelhalter, D. J. Best, N. G. Carlin, B. P. & Van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistics Society, 64*(Part 4), 583–639.

Train, K. E. (2003). *Discrete choice methods with simulation*. New York: Cambridge University Press.

Weerahandi, S. & Dalal, S. R. (1992). A choice-based approach to the diffusion of a service: forecasting fax penetration by market segments. *Marketing Science, 11*, 39–54.

Zeithaml, V. A., Bolton, R. N., Deighton, J., Keiningham, T. L., Lemon, K. N. & Petersen, J. A. (2006). Forward-looking focus: Can firms have adaptive foresight? *Journal of Service Research, 9*(2), 168–183.