

Reimer, Kerstin; Albers, Sönke

**Working Paper**

## Modeling Repeat Purchases in the Internet when RFM Captures Past Influence of Marketing

Arbeitspapiere des Instituts für Betriebswirtschaftslehre der Universität Kiel

Suggested Citation: Reimer, Kerstin; Albers, Sönke (2011) : Modeling Repeat Purchases in the Internet when RFM Captures Past Influence of Marketing, Arbeitspapiere des Instituts für Betriebswirtschaftslehre der Universität Kiel, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<http://hdl.handle.net/10419/50730>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Modeling Repeat Purchases in the Internet when RFM Captures Past Influence of Marketing

Kerstin Reimer<sup>1</sup> and Sönke Albers<sup>2</sup>

<sup>1</sup>Christian-Albrechts-University at Kiel  
Department of Innovation, New Media and Marketing  
Westring 425  
D-24098 Kiel  
Phone: +49(0)431 880-1552  
Fax: +49(0)431 880-1166  
Email: reimer@bwl.uni-kiel.de

<sup>2</sup>Kühne Logistics University  
Brooktorkai 25  
D-20457 Hamburg  
Phone: +49(0)40 328707-211  
Fax: +49(0)40 328707-209  
Email: soenke.albers@the-klu.org

Web: <http://www.bwl.uni-kiel.de/bwlinstitute/Innovation-Marketing/new/de/start/>  
<http://www.the-klu.org/faculty-and-research/resident-faculty/soenke-albers/>

## **Abstract**

Predicting online customer repeat purchase behavior by accounting for the marketing-mix plays an important role in a variety of empirical studies regarding individual customer relationship management. A number of sophisticated models have been developed for different forecasting purposes based on a – mostly linear – combination of purchase history, so called Recency-Frequency-Monetary Value (RFM)-variables and marketing variables. However, these studies focus on a high predictive validity rather than ensuring that their proposed models capture the original effects of marketing activities. Thus, they ignore an explicit relationship between the purchase history and marketing which leads to biased estimates in case these variables are correlated. This study develops a modeling framework for the prediction of repeat purchases that adequately combines purchase history data and marketing-mix information in order to determine the original impact of marketing. More specifically, we postulate that RFM already captures the effects of past marketing activities and the original marketing impact is represented by temporal changes from the purchase process. Our analysis highlights and confirms the importance of adequately modeling the relationship between RFM and marketing. In addition, the results show superiority of the proposed model compared to a model with a linear combination of RFM and marketing variables.

**Keywords** Repeat Purchase Forecasting Models, Marketing Actions, Generalized Bass Model, Media Downloads

## 1. Introduction

Getting a good idea of which model your customers' future purchase activities are following has been a major concern in both marketing research and practice since the Customer Relationship Management (CRM) concept swept the market in the 1990s.

In fact, there is significant amount of research on the analysis and prediction of customer purchasing behavior having produced a variety of models with high predictive validity for different forecasting purposes and product categories particularly in non-contractual online business settings (Gupta et al., 2006; Van den Poel & Buckinx, 2005). Generally, these forecasting models are in some way related to the well-known recency, frequency, and monetary value (RFM) framework which was introduced by Alden's catalog company in the 1920's (Roel 1988). For example, stochastic models for count data such as Pareto/NBD or BG/NBD<sup>1</sup> use recency and frequency information to predict the probability of a customer being active (Fader, Hardie, & Lee, 2005a, b; Kinshuk, Fader, & Hardie, 2010; Reinartz & Kumar, 2000, 2003). These models show high prediction performance with respect to forecasting individual customer repeat purchases and activity status, also called the probability of "being alive" or "p alive" which, in turn, can be used to estimate customer lifetime value (Reinartz & Kumar, 2000, 2003; Schmittlein & Peterson, 1994). More recently, modifications can be found such as the MBG/NBD model (Batislam, Denizel, & Filiztekin, 2007) explicitly incorporating zero purchasers or the Hierarchical Bayes extension to the Pareto/NBD to cope with today's customized marketing (Abe, 2009). Again, these sophisticated models perform very well in terms of the respective research objectives.

However, Wübben and von Wangenheim (2008) show that a model does not necessarily have to be sophisticated in order to precisely forecast a customer's transactions, especially with respect to managerial relevance and applicability. They prove that simple heuristics using

---

<sup>1</sup> Beta-geometric / negative binomial distribution

initial and repeat purchase data perform at least as well on the individual level as the stochastic models mentioned above. Similarly, Borle, Singh, and Jain (2008) find in their study on customer lifetime value measurement, that a simple heuristic approach based on average interpurchase time, average purchase amount, and average lifetime outperforms an extended Pareto/NBD model in terms of overall customer lifetime value (CLV) prediction.

Thus, it has been shown that the purchase history, often summarized as RFM, is a well-suited instrument for forecasting individual future behavior, not only with regard to sophisticated models but also to simple rules. Yet, from a marketing perspective, a prediction based on mere transaction data would clearly be insufficient as it misses important information on marketing activities. In order to optimally determine the impact of marketing instruments on purchasing behavior, these instruments must be adequately captured and combined with purchase history variables so that the resulting model

1. allows for a clear differentiation between the impact of RFM and the impact of marketing activities on repeat purchasing behavior,
2. has a higher predictive validity than a modeling framework with transactional data only and
3. provides plausible values for a subsequent optimization of the marketing-mix.

Considering, for example, the naïve option consisting simply of a linear combination of RFM and marketing variables we might observe the same phenomenon as Bass, Krishnan, and Jain (1994) did. The authors investigated the impact of marketing instruments on the diffusion process of new products using a Generalized Bass Model (GBM) with decision variables and found that the purchase history as reflected in the diffusion parameters already captures the process entirely. Thus, they conclude that the marketing effects cannot be estimated independently from the diffusion parameters. Given this evidence and further research on forecasting purchase behavior exemplarily described above, we believe that—analogueous to

Bass, Krishnan, and Jain (1994)—the marketing variables only shift the basic curve, i.e. the purchase process, temporarily, which in turn is determined, in our case, by RFM. In order to obtain this particular relationship, relative variables necessary for adequately modeling the temporal shifts have to be created similarly to the so-called mapping function suggested by Bass, Krishnan, and Jain (1994).

Therefore, we develop a repeat purchase forecasting model for panel count data on an individual level that multiplicatively connects RFM-variables, i.e. the interpurchase time or timing, purchase or transaction rate, and purchase amount or quantity, with marketing instruments as relative changes to the previous period. This formulation postulates that RFM-parameters shape the function according to our assumption while marketing variables, as “impulses”, only shift their projections; furthermore, it allows us to also capture the dynamics over time by reflecting the effects of lags in the projections (Bass, Krishnan, & Jain, 1994). Additionally, the model can serve as a basis for determining CLV and customer profitability as well as for optimizing marketing impact on purchase behavior (Borle, Singh, & Jain, 2008; Kumar & Shah, 2009; Reinartz, Thomas, & Kumar, 2005; Venkatesan, Kumar, & Bohling, 2007; Zhang & Wedel, 2009).

We show that with this particular modeling framework we are able to estimate marketing effects in a more adequate and more plausible way with coefficients representing elasticities and we are able to obtain a higher predictive validity.

We apply this framework to a rich customer dataset provided by a large online music download service with more than half a million customers. In fact, the emerging market for selling media products like movies, music, or audio books as downloads instead of offering to order them online has been particularly successful over the last few years and is still rapidly growing. As a consequence, music download services spend heavily on marketing activities in a variety of media to enhance customer retention and thus prevail against competitors and

piracy (IFPI Digital Music Report 2009). This very competitive and innovative environment, together with the presence of huge databases and detailed information on individual transaction histories and marketing activities makes it a highly relevant field for our proposed modeling framework.

In summary, this study contributes to existing research by developing a novel model for predicting customer repeat purchase behavior in an online environment. Moreover, we investigate its performance on a large dataset of a music download service, a new and important industry in the digital world, which is quite different from the typically analyzed consumer packaged goods (CPG) market (Reimer, Rutz, & Pauwels, 2010). To the best of our knowledge there is no such study that uses this particular nonlinear modeling framework for predicting customer repurchase behavior in the online business.

## **2. Background**

Understanding the impact of marketing activities on individual customer behavior is a key concept of customer relationship management and therefore a prerequisite to successfully implementing CRM systems. Extant research done in this area over the last decade mirrors the increasing importance of the analysis of advertising, direct marketing, and promotion effects in CRM contexts in a variety of different industries and business settings (Kumar & Shah, 2009; Manchanda et al., 2006; Rust & Verhoef, 2005; Venkatesan, Kumar, & Bohling, 2007). Particularly, with the breakthrough of new information and communication technologies, the understanding of customer purchase behavior with respect to marketing actions has risen to a new level: The availability of extensive individual customer data, predominantly in the online market, now allows for detailed analyses, thus helping managers and researchers to gain in-depth insight into customer purchasing processes and develop new marketing strategies. We are currently witnessing ever-increasing dynamics in the Internet (business) environment

requiring continuous empirical research on customer (repeat) purchase behavior by accounting for the marketing perspective (Abe 2009).

By briefly reviewing the literature regarding the impact of marketing instruments on online customer repurchase behavior, we find that previous studies use a variety of different methodologies and model applications. However, they are generally built on the same behaviorally-based purchase history variables (RFM framework) having been proven to be of strong explanatory power (Fader, Hardie, & Lee, 2005a).

In the following, we present an overview of different ways marketing variables are incorporated into repeat purchase forecasting models applied in the online business, and state that these modeling frameworks allow for parameter optimization with respect to marketing actions. Hereby, we differentiate between the purchase or transaction history representing pure transactional information such as RFM-variables, and past purchase behavior, which also incorporates information on carry-over effects and additional lagged variables to capture dynamic and periodical effects.

**One research stream** related to understanding individual customer purchase behavior including marketing information is represented by works of, for example, Ansari, Mela, and Neslin (2008), Lewis (2004), and Manchanda et al. (2006). These authors aim to identify relevant predictors for the repeat purchase probability by means of choice modeling such as logit, probit, or Tobit models, and survival models. These studies incorporate information on past purchases as well as marketing mix information, but only as a linear combination which could possibly produce the problems outlined in the introduction.

*Ansari, Mela, and Neslin (2008)*, for instance, develop a model of customer channel migration and apply it to a retailer operating over the Web and through catalogs selling consumer durable and apparel products in mature categories. They model purchase incidence and order-size components of purchase volume using a Type II Tobit specification and channel selection



using a probit model. In all specifications, the authors linearly connect marketing, i.e. emails and catalogs, and RFM-variables (Ansari, Mela, & Neslin, 2008, p. 62).

*Lewis (2004)* investigates the long-term effects of loyalty programs and other marketing instruments on customer retention at an Internet retailer specializing in nonperishable grocery and drugstore items by means of a discrete-choice dynamic programming formulation. He assumes that purchases represent the sequential choices of customers solving a dynamic optimization problem. The single-period utility functions determining the choice probabilities in the dynamic programming model are again linear functions of purchase history data and marketing variables such as email coupons (Lewis, 2004, p. 283 et seq.).

*Manchanda et al. (2006)* investigate the purchase behavior of customers who are exposed to banner advertising by the Web site of an Internet-only firm that sells health care/beauty products as well as nonprescription drugs. The authors model the potentially duration-dependent purchase incidence decision, i.e. whether and when to buy from the Web site, with a semi-parametric (constant piecewise) exponential hazard model in discrete time. More specifically, the model formulation focuses on the weekly purchase decision at the individual customer level, defined as a linear function of individual advertising exposure, browsing behavior variables, timing of the last purchase and unobserved heterogeneity. The modeling framework is analogous to purchase incidence models, such as the binary logit/probit with temporal fixed effects (Manchanda et al., 2006, 101 et seq.). Thus, these authors also linearly connect utility from marketing with purchase (and browsing) history information, which could ultimately lead to the same potential problem observed in the previous studies: the extraction of the original marketing effects when marketing is absorbed by RFM, i.e. the variables are partially redundant. Consequently, the models of this research area do not provide a relevant framework for our analysis.

**A second stream of research** includes studies on maximizing CLV and customer profitability by incorporating marketing-mix information in addition to RFM-variables into the purchase forecasting model such as Rust and Verhoef (2005), Venkatesan and Kumar (2004), and Venkatesan, Kumar, and Bohling (2007), which are discussed below.

*Rust and Verhoef (2005)* provide a model for optimizing multiple marketing interventions in intermediate-term customer relationship management (CRM) using data from a financial service provider. To optimize individual customer profitability, they model the impact of marketing interventions (direct mailings and relationship magazines) on customers' changes in gross profits. Moreover, the authors propose a hierarchical model to account for customer heterogeneity at the customer level. The first level is defined as the general shift in gross profit model which is a linear function of the marketing intervention levels targeted at customer  $i$ . The second level contains the customer-specific marketing response parameter vector (from level one) which is regressed on purchase history variables and customer characteristics of customer  $i$ . However, when using a hierarchical linear model, it is possible that the variables at the second level, i.e. the within group control variables such as the purchase history variables, can offset the variables at the first level especially when marketing is significantly correlated with RFM. In this case, instable and/or inconsistent estimates would be the result. This, in turn, would also lead to biased optimization results for the marketing interventions. Therefore, we do not consider this modeling framework for our research purpose either.

*Venkatesan and Kumar (2004)* present a framework to predict CLV for customer selection based on data from a large multinational computer hardware and software manufacturer. They incorporate a stochastic model to predict each customer's purchase frequency combined with a latent class segmentation model to account for heterogeneity. More specifically, the authors use the generalized gamma model of interpurchase timing developed by Allenby, Leone, and

Jen (1999), which they multiply with the probability that a purchase occasion  $j$  from a customer  $i$  belongs to subgroup  $k$  based on a cumulative normal distribution (probit). The probit is a function of the marketing decision variables consisting of levels of rich and standard modes of communication and intercontact time, and of covariates such as bidirectional communication, cross buying and upgrading activity (Venkatesan & Kumar, 2004, p. 113).

A more recent study by *Venkatesan, Kumar, and Bohling (2007)* extends the CLV framework of Venkatesan and Kumar (2004) by developing a joint model, implementing the same data set, to simultaneously predict a customer's future purchase timing and purchase quantity, likewise assuming that the population consists of  $k$  subgroups. The authors define a probability model for timing and quantity based on the same concomitant mixture framework (Allenby, Leone, & Jen, 1999) and the same probit function Venkatesan and Kumar (2004) used to determine the segment probability. The purchase rate parameter is modeled as a function of past purchase behavior to capture the effect of the level of previous purchases and the frequency of previous purchases on the timing of the current purchase occasion (Venkatesan, Kumar, & Bohling, 2007, p. 586).

In both studies, the density functions for timing as well as for timing and quantity respectively include information on the purchase history; the density function for quantity in the latter study also contains covariates. Marketing decision variables (in absolute values) are only incorporated in the segmentation procedure but implicitly influence purchase frequency (and quantity) because the segment probabilities serve as weights for the prediction of the interpurchase time of each subgroup. That is, by applying these frameworks developed for customer selection using CLV the authors separate the effects of marketing and past purchase data; i.e., they include them into different sub-functions of the total model, yet without imposing an explicit relationship between marketing and RFM-variables. More specifically,

in case there is a direct link between them as we postulate, i.e. when the purchase history captures past influence of marketing, both sub-models partially include the impact of marketing and it would not be clear how to extract the original/actual effect of these decision variables.

**A third research stream** concentrates on the effects and optimization of promotional activity on customer repurchase behavior in the online (and online versus offline) environment (Khan, Lewis, & Singh, 2009; Zhang & Krishnamurthi, 2004; Zhang & Wedel, 2009).

*Khan, Lewis, and Singh (2009)* formulate a joint (dynamic programming) model of timing and expenditure in a discrete-choice hazard modeling framework investigating and optimizing the effectiveness of different individual-level promotions at an online retailer selling nonperishable grocery and drugstore items. The authors define the individual utility function for making a purchase in period  $t$  as a linear combination of covariates with individual-specific time-varying coefficients such as price, presence of a coupon, and shipping fees, and covariates with time-invariant coefficients, e.g. status in loyalty program and demographics. Additionally, the former are multiplied by a function capturing the underlying duration dependence of the response parameters. That function mainly includes the time since the last purchase and frequency. To model expenditure conditional on purchase incidence, they use a semi-log specification and the same covariates with individual-specific time-varying coefficients as for the incidence utility plus two purchase history variables (Khan, Lewis, & Singh, 2009, p. 1067). As a conclusion, Khan, Lewis, and Singh (2009) explicitly investigate the relationship between promotional and purchase history variables in the utility for purchase incidence by also estimating their interactions in addition to the main effects. However, the objective of this operationalization is to measure how the effects of promotions (in absolute values) vary with recency and frequency of purchase whereas we aim to develop a model that adequately captures the effects of marketing in combination with the purchase history.

Moreover, our model allows for parameter optimization of the marketing-mix by introducing relative variables enabling us to additionally determine the optimal level of advertising and promotional activity.

*Zhang and Krishnamurthi (2004)* and *Zhang and Wedel (2009)* investigate and optimize the effectiveness of customized price promotions (different price cuts) both using similar modeling frameworks. The former provide a decision-support system of customized promotions in online stores which are not only tailored to individual households but also dynamically updated on each shopping occasion. Zhang and Krishnamurthi (2004) develop a joint (simultaneously estimated) purchase, incidence-brand choice-purchase quantity model that allows for the variety-seeking/inertia behavior to differ among households and change within a household over time, using data for butter and liquid detergent sold at a leading online grocery retailer.

By again focusing on the functional forms, we find that the joint probability of purchase incidence and choice is formulated as a nested logit model with two utility functions: a brand utility and a threshold utility for purchase incidence. The brand utility of alternative  $k$  for a household  $i$  during a shopping trip  $t$  is a function of brand-specific constants, marketing-mix variables such as regular price and price discount, and a time-varying purchase event feedback effect component determining inertia/variety seeking pattern. The threshold function only incorporates purchase variables (Zhang & Krishnamurthi, 2004, p. 562). The quantity for alternative  $k$  and household  $i$  conditional on purchase incidence and brand choice is determined by a linear function of the household  $i$ 's average purchase quantity, purchase frequency, and alternative  $k$ 's regular price and price cut.

Zhang and Wedel (2009) also examine the effectiveness, i.e. the profit potential in particular, of customized price promotions but this time at three levels of granularity (mass market, segment and individual specific), and in online and offline stores. They use the same data of

the Internet grocery retailer mentioned in the previous study and additional data of the offline retailer through which the online retailer procures its products in order to create a comparable base. The authors build their model on the same joint purchase incidence-brand choice-purchase quantity model developed by Zhang and Krishnamurthi (2004) and extend it by adding the feature advertising/display to the model specification for the offline data set.

In summary, Zhang and Krishnamurthi (2004) and Zhang and Wedel (2009) nicely combine purchase incidence and brand choice in a nested logit model in which information on price promotions is used to determine the brand utility together with predictors accounting for the impact of past purchases on current brand preference. The category threshold representing the baseline utility for purchase incidence is solely built on traditional purchase history variables (frequency, monetary value). Thus, they postulate that marketing activities, i.e. price cuts, primarily affect brand choice decision given that the threshold utility based on transactional data has been exceeded by the utilities of the alternatives under consideration. This is a common model formulation when investigating the consumer packaged goods (CPG) market where many brands and various (promotional) prices exist for the same product, such as detergent or butter, and to account for inventory effects influencing purchase timing.

However, the purchase situation in our market of interest—the music download industry—is quite different than the one in an (online) supermarket. First of all, customers do not have to make a brand choice decision as every piece (song, album) is unique. Moreover, we generally do not find inventory effects due to the nature of the product: there is no such need for the product like for detergent. Music is a non-utilitarian/hedonic product, and it does not disappear after consumption. These products are purchased rather spontaneously, in contrast to the planned shopping “trip” for CPGs; thus, purchase incidence in a music download setting should also be influenced by marketing actions and not solely depend on the purchase history. Finally, music downloads benefit from instantaneous delivery and can be consumed

right after purchase whereas online orders for CPGs need to be processed and delivered offline.

Based on this reasoning, we believe that the modeling frameworks of Zhang and Krishnamurthi (2004) and Zhang and Wedel (2009) cannot be adequately applied to our data and research objective. Indeed, the fact that music downloads are significantly different from CPGs should be kept in mind with respect to all studies presented in this section. It has specifically been outlined in connection with these last two studies because they include the brand choice decision in addition to timing and quantity.

We aim to add to this research by proposing a model that imposes an explicit—and other than a linear—relationship between the purchase history and marketing impact, and adequately represents purchase behavior of online media downloads focusing on digital hedonic products such as music, movies, or ebooks.

In fact, we postulate that the effects of marketing activities are already absorbed by past purchase behavior, which requires an explicit functional form including marketing as relative changes over time. Additionally, our framework should allow for parameter optimization. The modeling approach and the motivation for choosing that particular framework will be explained in detail in the following section.

### **3. Modeling Approach**

We develop our model using the same idea as in the Generalized Bass Model (Bass, Krishnan, & Jain, 1994) providing the following favorable properties important for our proposed model.

It has been found that, over a large number of new products, the Bass model describes the empirical adoption curve quite well, which in our case represents the purchase history of each customer. Bass, Krishnan, and Jain (1994) extend the base model to the so-called Generalized Bass Model (GBM) with decision variables and show that the Bass model always fits well

even without decision (marketing) variables. They postulate that under normal market conditions, marketing can only temporarily shift the basic diffusion function. For example, if—assuming the GBM to be the true model—marketing and diffusion process variables are highly correlated, e.g. marketing changes by more or less a constant in each period, we will also get a good fit to the data by just using the regular Bass model though, technically, it is a misspecification. However, if marketing variables are statistically significant, the GBM provides a better fit explaining the deviations from the curve of the Bass model.

Given the favorable features of this modeling framework, we build our purchase forecasting model analogous to the idea of the GBM: we assume that RFM-variables shape the function providing a smooth curve while marketing variables only shift their projections. That is, we use relative changes of marketing variables instead of absolute values to model the impulses and multiplicatively connect them to the purchase history data in order to link the impulses to the purchase process thereby avoiding any correlation between RFM and marketing variables (like in the Bass model).

Moreover, we conceptualize the consumer repurchase decision to be consisting of multi-stages, similarly to Andrews and Currim (2009). In absence of a possible brand choice decision, as each product of our relevant market is unique, we model two stages: purchase incidence and purchase quantity. We particularly focus on the incidence decision for which we develop a specific modeling framework using the idea of the Generalized Bass Model (Bass, Krishnan, & Jain, 1994).

### *3.1 Base model*

In this section, we discuss the base specification of the purchase incidence and the purchase quantity model. Starting with the purchase incidence model, the probability that a customer i



decides to make a purchase at time t is represented by a logit model, here applied to panel data (Maddala, 1987):

$$P_{it}(\textit{incidence}) = \frac{\exp(U_{it})}{1 + \exp(U_{it})}. \quad (1)$$

The utility that customer i obtains from making the purchase at time t is given by the following non-linear function:

$$U_{it} = (\alpha_0 + \boldsymbol{\alpha}'\mathbf{X}_{it}) \cdot (1 + \boldsymbol{\beta}'\mathbf{Z}_t) \quad (2)$$

where  $X$  is a matrix of time-varying customer individual purchase history variables,  $Z$  a matrix of the positive relative changes ( $\Delta$ ) of different marketing instruments also varying over time,  $\alpha_0$  a random individual-specific intercept term and  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\beta}$  the respective vectors of coefficients. The positive relative changes are calculated using the formula:  $[\max(0, \Delta \text{marketing}(t))] / \text{marketing}(t-1)$  suggested by Simon (1982), who argued that changes in marketing should have an impact only if they are positive, thus resulting in asymmetric effects of marketing change. Bass, Krishnan, and Jain (1994) have already taken the same operationalization for calculating the percentage changes of the decision variables in their GBM.

Given that formulation (Eq. 2), we posit—as explained above—that the RFM-part captures the whole purchase history of each customer which we define as the base curve analogous to the diffusion curve in the Generalized Bass Model. By using the relative changes of marketing decision variables which only shift the projections, i.e. the forecast from RFM, we assume that reactions to past marketing actions and other lagged effects are already reflected in the projections. Thus, the impact of marketing actions at time t on repeat purchasing behavior can clearly be differentiated from the explanatory power of the purchase history by simultaneously avoiding potential collinearity between RFM and marketing as well.

For the second stage, the quantity model, we define a negative binomial panel regression model conditional on the incidence decision by using the prediction of the first stage as an additional explanatory variable similar to the two-step Heckman estimation in the cross-sectional case (Heckman, 1979; Puhani, 2000). We estimate the two stages separately because estimating a logit and a negative binomial panel regression model simultaneously using the particular functional form in the logit model presented above and applying it to a very rich data set is infeasible due to the size of the likelihood and the resulting computing time.

The standard negative binomial (regression) model for count data is an appropriate framework for explaining and forecasting purchase behavior with respect to the number of products bought and can be found in a variety of applications in marketing —albeit often without explanatory variables (Fader, Hardie, & Lee, 2005a; Batislam, Denizel, & Filiztekin, 2007). The negative binomial distribution (NBD) is a generalization of the Poisson distribution, and implies a Poisson purchase process at the individual-level with purchase rate parameter  $\mu_{it}$  following a Gamma distribution. Hence, the probability distribution for the NBD—in our case panel regression—with  $y_{it}$  as the number of independent events that occur during a fixed time period is as follows:

$$P(y_{it} | \mathbf{x}_{it}, \theta) = \frac{\Gamma(y_{it} + \theta)}{\Gamma(\theta) y_{it}!} \left( \frac{\theta}{\theta + \mu_{it}} \right)^\theta \left( \frac{\mu_{it}}{\theta + \mu_{it}} \right)^{y_{it}} \quad y_{it} = 0, 1, 2, \dots \quad (3)$$

where  $\theta$  is the systematic (or overdispersion) parameter of the Gamma distribution,

$$E(y_{it} | \mathbf{x}_{it}) = \mu_{it} = e^{\mathbf{x}_{it}' \boldsymbol{\gamma}} \quad (4)$$

the expected value, and

$$V(y_{it} | \mathbf{x}_{it}) = \mu_{it} + (\mu_{it}^2 / \theta) \quad (5)$$

the conditional variance of  $y_{it}$  with  $x_{it}$  as the vector of explanatory variables which includes the prediction from the logit model, purchase history variables as well as marketing

information.  $\gamma$  represents the vector of coefficients. Allowing for gamma heterogeneity in the purchase rate is particularly useful when dealing with overdispersed count data, which is typically the case in panel data settings with small time units such as weeks (Abe, 2009).

### 3.2 *Model Specification*

We specify the regression equation for the utility in the logit panel model for purchase incidence as follows: The purchase history defined in the first factor is represented by the three well-known variables recency (rec), frequency (freq), and monetary value (mv). Moreover, we include the quadratic term of frequency (freq<sup>2</sup>) postulating that the relationship between frequency and purchase incidence is inverse-u shaped according to, for example, Venkatesan, Kumar, and Bohling (2007). The second factor of the model incorporates a linear function of the most common offline and online advertising media, i.e. TV (TV), print (P) and Internet (I) banner advertising, in addition to the widely-used coupon promotions (C), all expressed in positive relative changes. Whereas the advertising instruments are exactly calculated according to Simon (1982), we had to employ a slight modification for determining the relative changes of coupon promotions being a dummy variable as an indicator for promotional activity in period t. Instead of just dividing by marketing (t-1), we divided by the mean of marketing (t) and marketing (t-1) to achieve in any case a positive quotient which would otherwise always be zero. Finally, we linearly added a trend and diverse covariates to control for seasonality, observed heterogeneity and other exogenous shocks.

In order to reduce the high computational burden resulting from the estimation of a non-linear utility function in a logit panel model, we linearized that function by multiplying the two factors out, leading to interaction terms which need to be estimated jointly. The utility equation and its transformation for a customer i at time t are given below:

$$U_{it} = (\alpha_0 + \alpha_1 rec_{i,t-1} + \alpha_2 freq_{i,t-1} + \alpha_3 freq_{i,t-1}^2 + \alpha_4 mv_{i,t-1}) \cdot (1 + \beta_1 \cdot (\Delta TV_t / TV_{t-1}) + \beta_2 \cdot (\Delta R_t / R_{t-1}) + \beta_3 \cdot (\Delta I_t / I_{t-1}) + \beta_4 \cdot (\Delta C_t / ((C_t + C_{t-1}) / 2))) + \text{covariates} =$$

$$\begin{aligned} & \alpha_0 + \alpha_1 \cdot rec_{i,t-1} + \alpha_2 \cdot freq_{i,t-1} + \alpha_3 \cdot freq_{i,t-1}^2 + \alpha_4 \cdot mv_{i,t-1} \\ & + \alpha_0 \cdot \beta_1 \cdot (\Delta TV_t / TV_{t-1}) + \alpha_0 \cdot \beta_2 \cdot (\Delta R_t / R_{t-1}) + \alpha_0 \cdot \beta_3 \cdot (\Delta I_t / I_{t-1}) + \alpha_0 \cdot \beta_4 \cdot (\Delta C_t / ((C_t + C_{t-1}) / 2)) \\ & + \alpha_1 \cdot \beta_1 \cdot rec_{i,t-1} \cdot (\Delta TV_t / TV_{t-1}) + \alpha_1 \cdot \beta_2 \cdot rec_{i,t-1} \cdot (\Delta R_t / R_{t-1}) + \alpha_1 \cdot \beta_3 \cdot rec_{i,t-1} \cdot (\Delta I_t / I_{t-1}) \\ & + \alpha_1 \cdot \beta_4 \cdot rec_{i,t-1} \cdot (\Delta C_t / ((C_t + C_{t-1}) / 2)) \\ & + \alpha_2 \cdot \beta_1 \cdot freq_{i,t-1} \cdot (\Delta TV_t / TV_{t-1}) + \alpha_2 \cdot \beta_2 \cdot freq_{i,t-1} \cdot (\Delta R_t / R_{t-1}) + \alpha_2 \cdot \beta_3 \cdot freq_{i,t-1} \cdot (\Delta I_t / I_{t-1}) \quad (6) \\ & + \alpha_2 \cdot \beta_4 \cdot freq_{i,t-1} \cdot (\Delta C_t / ((C_t + C_{t-1}) / 2)) \\ & + \alpha_3 \cdot \beta_1 \cdot freq_{i,t-1}^2 \cdot (\Delta TV_t / TV_{t-1}) + \alpha_3 \cdot \beta_2 \cdot freq_{i,t-1}^2 \cdot (\Delta R_t / R_{t-1}) + \alpha_3 \cdot \beta_3 \cdot freq_{i,t-1}^2 \cdot (\Delta I_t / I_{t-1}) \\ & + \alpha_3 \cdot \beta_4 \cdot freq_{i,t-1}^2 \cdot (\Delta C_t / ((C_t + C_{t-1}) / 2)) \\ & + \alpha_4 \cdot \beta_1 \cdot mv_{i,t-1} \cdot (\Delta TV_t / TV_{t-1}) + \alpha_4 \cdot \beta_2 \cdot mv_{i,t-1} \cdot (\Delta R_t / R_{t-1}) + \alpha_4 \cdot \beta_3 \cdot mv_{i,t-1} \cdot (\Delta I_t / I_{t-1}) \\ & + \alpha_4 \cdot \beta_4 \cdot mv_{i,t-1} \cdot (\Delta C_t / ((C_t + C_{t-1}) / 2)) \\ & + \delta \cdot trend + \text{covariates} + \varepsilon_{it} \end{aligned}$$

where  $\alpha_j$  are the coefficients of the purchase history variables,  $\beta_k$  the coefficients for the marketing decision variables,  $\delta$  the trend coefficient and  $\varepsilon_{it}$  the logit error.

The regression equation of the negative binomial model is specified as follows:

$$\begin{aligned} \mu_{it} = & \exp(\gamma_0 + \gamma_1 \cdot \text{logit\_predict}_{it} + \gamma_2 \cdot mv_{i,t-1} + \gamma_3 \cdot TV_t + \gamma_4 \cdot R_t + \gamma_5 \cdot I_t + \gamma_6 \cdot C_t \\ & + \delta \cdot trend + \text{covariates} + \eta_{it}) \end{aligned} \quad (7)$$

with  $\text{logit\_predict}$  representing the prediction from the logit model,  $\gamma_j$  as the coefficients of the explanatory variables and  $\eta_{it}$  as the individual error term. Apart from the prediction of the incidence model particularly reflecting recency and frequency information, we believe that the cumulative revenue up to t-1 (MV) and the marketing instruments significantly influence the quantity a customer purchases, which is in line with existing research (Zhang & Wedel, 2009; Khan, Lewis, & Singh, 2009).

### 3.3 Estimation

Before estimating the proposed modeling framework, two preliminary analyses must be performed. First of all, we calculate a carry-over for each advertising variable measured in Gross Rating Points (GRPs) based on the grid stock search model (Greene, 2003, 566 et seq.), following common practice of extant previous research (e.g. Raman & Naik, 2006).

Second, in order to prove our assumption that the RFM-variables do not only capture the whole purchase history but also absorb the impact of past marketing activities, we estimate a Seemingly Unrelated Regression (SUR) model, which qualifies as an adequate method for the analysis of our interest. SUR models have a number of correlated regression equations and allow each of the dependent variables to have a different design matrix with some or all of the predictor variables being the same (Greene, 2003, 340 et seq.). Applied to our specific case, we regress each of the four purchase history variables on the advertising instruments, which include a carry-over and on the coupon promotion variable.

Finally, we estimate the linearized proposed purchase incidence model (Eq. 6) and three benchmark models as listed in Table 1, chosen according to our reasoning: Model 2 should outperform model 1 in case we find an inverse-u shaped relationship between purchase incidence and frequency. Model 3 is assumed to be superior to model 2 if marketing decision variables have a significant impact on purchase incidence albeit being potentially highly correlated with RFM-variables. The proposed model contains the same set of variables, but in contrast to model 3 clearly separates the impact of marketing and RFM and allows for optimization by implementing the discussed functional form. We compare the performance of all models based on the Bayesian Information Criterion (BIC) and McFadden  $R^2$ , which are widely used metrics in the marketing literature.

**Insert Table 1 Here**

Due to the transformation of the proposed incidence model into a linear function the majority of the resulting coefficients are composed of interaction terms. Given only one equation, it is impossible to solve for the individual coefficients. Therefore, we insert the values of the RFM-variables of each customer and point in time and calculate individual time-varying coefficients (elasticities). To be able to evaluate the results, i.e. the elasticities for marketing decision variables in particular, we determine the weighted mean of each coefficient over all customers and the total observation period. Details on the inference procedure are given in the results section.

The purchase quantity model is estimated as defined in Equation 7. For both stages, we chose a random effects panel model framework for several reasons: We deal with a large number of cross-sectional units and also incorporate time-invariant observations plus we want to make inferences about the population these cross-sectional data come from (Maddala, 1987). Moreover, we believe that a random effects model specification is an adequate method for evaluating the performance of our proposed model and different benchmark models with respect to our research objective. It produces robust results and, as opposed to a random coefficients modeling framework, it does not lead to the non-necessary enormous computational complexity we would have to face by estimating the above described functions using customer-individual coefficients. We rely on the evidence from previous research that simple models perform quite well compared to sophisticated models (Fader & Hardie, 2005; Wübben & von Wangenheim, 2008).

## **4. Empirical Application**

### *4.1 Data*

Our data represent the whole customer database of a major music download company with more than half a million customers over a time period of 20 months (87 weeks) starting in

January 2005, and includes rich information on an individual level as well as on an aggregate level. We underline again that music downloads as digital hedonic products have very specific characteristics fairly different from the usually investigated CPG or durable products which needs to be taken into account when performing the analyses and interpreting the results. (Details have been given in the literature background section). With respect to holdout validation, we use 79 weeks to calibrate the model, and 8 weeks to investigate the forecasting performance of customer purchase behavior representing a realistic short-term planning horizon. However, estimating a random effects panel model with a high number of explanatory variables using all customers is infeasible due to the size of the likelihood and the resulting computing time. Therefore, we draw a random sample of 5,478 (~1%) of all customers being active<sup>2</sup> during the respective period.

We focus our study on existing customers only in order to clearly identify the drivers of online repeat purchase behavior, separate from the customer acquisition process. By using a random sample of the total customer database, the data also includes newly acquired customers within the observation period, i.e. first-time buyers, of which we eliminate the first purchase in order to adequately calculate the recency and frequency, i.e. the repeat purchase variables of each customer and hence obtain the relevant data for our analysis (Venkatesan & Kumar, 2004).

The data contains weekly information on individual customer spending, i.e., the number of items purchased and the revenue per week, as well as information on various marketing actions and some customer characteristics. Our dependent variable, the number of items bought in a week, ranges from 0 to 122 (see Table 1). Customers typically buy more than one item per transaction, and mostly do not purchase more than once a week. A proxy of the

---

<sup>2</sup> „Being active“ includes each customer who made at least two transactions in the 87 weeks.

individual revenue per week can be obtained by multiplying the number of items with the average product price since prices differ only marginally across music downloads.

Based on the transaction data we also develop the well-known RFM-variables, recency, frequency and monetary value, by maintaining the panel data framework. This means that, each of the three predictors varies over time. The individual frequency in week  $t$  (or weekly transaction rate) of customer  $i$  equals 1 if he or she has made at least one purchase in week  $t-1$ ,<sup>3</sup> and decreases continuously until the next purchase is made, calculated as 1 divided by the number of weeks since the last purchase. In a cross-sectional analysis, the recency usually represents the time since the last purchase, in this case measured as number of weeks. However, given our panel model structure the customer-individual recency is also a time-varying variable and, by definition, highly correlated with the above-described frequency as it simply represents the reciprocal value of the frequency. Being essential for our analysis, we replace the recency by its mean-centered values; i.e., we calculate the mean recency for customer  $i$  over the 87 weeks and subtract this mean from the actual recency of customer  $i$  in week  $t$ . This operationalization has already been used by Chintagunta and Haldar (1998), but with respect to purchase quantity in order to avoid endogeneity issues in their model. Thus, negative values imply a short(er) time since the last purchase; conversely, values close to zero and positive values a longer time. The last RFM-variable, the monetary value in week  $t$ , is defined as the cumulative revenue of customer  $i$  up to week  $t-1$ .

The high standard deviations which we find for all purchase history variables reveal that the customer purchasing behavior is very, thus heterogeneous supporting the application of individual-level forecasting models.

---

<sup>3</sup> An examination of the data revealed that 98% of the 346,882 observations did not purchase multiple times in any given week so that we can consider our model with this unit of observation as appropriate.



### **Insert Table 2 Here**

In addition to the RFM-variables, which are assumed to capture the whole purchase history, we include different marketing actions of which we have information on a weekly basis. In particular, we investigate the effects of TV and radio advertising which are measured in gross rating points (GRP) as well as the impact of internet advertising in the form of banner ads available as the number of days per week it is present. As a complement to the advertising data, we also have information on coupon actions over the observation period included as a dummy variable. Like in most of the studies, these variables are only available on an aggregate market level.

In terms of frequency, TV is the advertising instrument the firm uses most often, with 57 out of 87 weeks. However, when it comes to volume significant differences emerge. TV has a weekly average of 36 GRPs compared to radio advertising which the firm used only sparsely with just four radio campaigns over a total of nine weeks. Yet, the biggest radio campaign that lasted three weeks has a comparatively high exposure-level with 162 GRPs per week. Internet (banner) advertising increases over the observation period—from 14 weeks in 2005 to 32 weeks in 2006. As described in the modeling approach, we only use the positive relative changes, i.e. the “impulses”, of all advertising variables after a carry-over effect has been included. We calculated the carry-over (based on the grid stock search model (Greene, 2003: 566 et seq.)) using weekly aggregated sales as dependent variable resulting in carry-over values of .90 for TV, of .78 for radio and .88 for Internet. These results confirm prior findings (Naik & Raman, 2003; Naik, Raman, & Srinivasan, 2007).

### **Insert Table 3 Here**

We also study the effects of newsletter and permission mailings as a direct marketing activity for which customers have to sign-up, and if the first purchase was made by using a coupon

promotion. Both variables are measured as binary variables and are sent to approximately 20% of all customers. Emails with newsletters are sent out every week, whereas permission mailings are sent irregularly related to special events or holidays.

The data also include information that enables us to control for seasonality as well as for new releases of famous artists and bands. Finally, gender, as available demographic variable, is incorporated and an additional transaction-based variable, which represents the number of periods between registration and first purchase. We assume that the longer this period, the higher the probability that this customer will make repeat purchases; we consider that first purchase happening significant time after the initial registration as commitment or active decision for that particular music download service over competitors.

#### *4.2 Estimation results*

First of all, we present the results of the SUR model shown in Table 4a and 4b, estimated to test the relationship between marketing and RFM-variables. All variables are highly significant on a 1%-level, proving our assumption that the marketing instruments are significantly correlated with the purchase history variables.

#### **Insert Table 4a Here**

Moreover, the result of the Breusch-Pagan test of independence based on the correlation matrix of the residuals reveals that we can reject the null hypothesis of equal residuals (see table 4a). This means that a SUR model is preferred over separate OLS regressions, which do not produce efficient estimates in this case. The positive estimates for advertising in the recency regression (value in  $t$  represents mean-centered recency of  $t-1$ ) and the negative advertising estimates in the frequency regressions (value in  $t$  represents frequency of  $t-1$ )

indicate that advertising effort is high in times of lower purchase activity. Concerning the regression with monetary value as the dependent variable we find a positive relationship between the advertising instruments and the monetary value whereas coupons are negatively related to that variable.

**Insert Table 4b Here**

Next, we estimated the different benchmark logit models for the incidence model outlined in section 3.3. Table 5 shows their performance evaluated by the log-likelihood (LL), the Bayesian Information Criterion (BIC) and the pseudo McFadden  $R^2$ . The statistics reveal the hypothesized relative performance with model 1 performing notably worst confirming the inverse u-shaped relationship between frequency and purchase incidence implemented in model 2. Model 3 performs only slightly better than the model without marketing, which proves our assumption – analogous to Bass, Krishnan, and Jain (1994) – that RFM already captures the purchase process very well. Our proposed model performs best, indicating that marketing variables do indeed have a significant impact on music download behavior, even though the difference of the fit to model 3 with the linear combination of both variable sets is rather small. Despite this small difference, we find significant effects of the marketing instruments, which implies that the effect of the RFM-variables is biased as long as marketing effects are omitted. In fact, besides the improvement itself, it is important to assess if the proposed model also produces (more) plausible coefficients owing to the particular structure which – to our opinion – provides an adequate combination of RFM and marketing. In the following, we will discuss the model coefficients in detail. Due to superiority and space limitation, we only present the results of the proposed model, which applies to the quantity model as well.

### **Insert Table 5 Here**

Table 6 shows the coefficients and standard errors of all variables from the logit panel estimation (incidence model) listed according to the type of variable. The constant and the RFM-variables – all highly significant - are given first. The mean-centered recency has a positive effect, which can be interpreted as ‘the longer the interpurchase time the higher the probability of a repurchase’, which is in line with previous literature (Ansari, Mela, & Neslin, 2008; Khan, Lewis, & Singh, 2009). The signs of frequency (+) and square of frequency (-) reveal that frequency has an inverse u-shaped impact on the purchase probability. More specifically, the propensity for a purchase increases for a certain time after a purchase event up to a maximum (reaching the customer’s optimal frequency), and then decreases together with the decreasing frequency, i.e. a very low frequency has a negative effect on the purchase probability (Reinartz, Thomas, & Kumar, 2005). The monetary value shows the expected positive impact on purchase incidence, which is consistent with existing research and can be interpreted as an indicator for customer retention: the more music a customer has already purchased, the higher the probability of a repurchase (Kumar & Shah, 2009; Lewis, 2004).

### **Insert Table 6 Here**

The coefficients of the marketing variables and the interaction terms cannot be interpreted in the form displayed in table 6 because each of them represents an interaction of two coefficients as a result of the linearized function (see Eq. 6). Therefore, we propose an approach described below table 6 – we call it ‘inference procedure’ – to extract the individual coefficients for TV, radio, and Internet advertising, as well as coupon promotions.

With respect to the covariates, several interesting results can be found. The coefficient of the customer-individual trend (starting at the first (second) purchase of each customer) has a negative sign, significant on the 1%-level, revealing that customers tend to be more active in

the beginning, i.e. when they are new customers.<sup>4</sup> Gender does not play a significant role nor does newsletter or permission. In case a customer uses a coupon with the first purchase we find evidence displayed in the significant negative coefficient that these customers primarily want to benefit from the promotion, and generally do not intend to further use this music download provider. However, customers whose registration and first purchase happen to be at separate points in time tend to be more committed, as demonstrated by the significant, positive coefficient of “Time from registration to 1<sup>st</sup> purchase”. This may be explained by the active decision process preceding the purchase since the customer could just as well sign up for a different service with nearly the same effort and experience and make a purchase there. Moreover, a release of a new single by one of the Top100 artists or bands has a positive impact on the purchase probability confirming the relevance of controlling for market activity in addition to marketing information.

Advertising sometimes features new releases, but it is necessary to capture these effects separately. The correlation between new releases and advertising is positive but moderate (.08-.15). The remaining variables control for seasonality and holidays, and predominantly show significant effects revealing that the music download industry is strongly influenced by seasonality.

### *Inference procedure*

The following procedure must be performed for each advertising coefficient separately. We will explain the steps by focusing on TV advertising, i.e. we only use the relevant parts for TV from the linear formulation. Under this condition, Equation 6 can be rewritten as:

---

<sup>4</sup> Only 5% of all customers already existed before the observation period. For the other 95% we eliminated the first purchase in order to separate acquisition from repeat purchasing behavior (see also 4.1 data).

$$\begin{aligned}
& (\alpha_0 + \alpha_1 rec_{i,t-1} + \alpha_2 freq_{i,t-1} + \alpha_3 freq_{i,t-1}^2 + \alpha_4 mv_{i,t-1}) \cdot \beta_1 \cdot (\Delta TV_t / TV_{t-1}) = \\
& (\alpha_0 \cdot \beta_1 + \alpha_1 \cdot \beta_1 \cdot rec_{i,t-1} + \alpha_2 \cdot \beta_1 \cdot freq_{i,t-1} + \alpha_3 \cdot \beta_1 \cdot freq_{i,t-1}^2 + \alpha_4 \cdot \beta_1 \cdot mv_{i,t-1}) \cdot (\Delta TV_t / TV_{t-1}) \quad (8)
\end{aligned}$$

and solved for  $\beta_1$ :

$$\Rightarrow \beta_1 = \frac{(\alpha_0 \beta_1 + \alpha_1 \beta_1 \cdot rec_{i,t-1} + \alpha_2 \beta_1 \cdot freq_{i,t-1} + \alpha_3 \beta_1 \cdot freq_{i,t-1}^2 + \alpha_4 \beta_1 \cdot mv_{i,t-1})}{(\alpha_0 + \alpha_1 rec_{i,t-1} + \alpha_2 freq_{i,t-1} + \alpha_3 freq_{i,t-1}^2 + \alpha_4 mv_{i,t-1})} \quad (9)$$

where the coefficients  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ ,  $\alpha_4$  and interactions  $\alpha_0\beta_1$ ,  $\alpha_1\beta_1$ ,  $\alpha_2\beta_1$ ,  $\alpha_3\beta_1$ ,  $\alpha_4\beta_1$  are known from the logit panel model estimation (see Table 6). The purchase history variables are then replaced by their individual values varying over time and customer so that individual  $\beta_{1,ii}$  could be calculated. Finally, we aggregate these values to a weighted mean  $\bar{\beta}_1$  representing the current effects' elasticity of TV advertising. The coefficients for  $\beta_2$ ,  $\beta_3$  and  $\beta_4$  are determined analogously. The standard errors of  $\beta_i$  are calculated according to

$$std.err(\beta) = \frac{\sqrt{\sum (y - \hat{y})^2 / (n - k - 1)}}{\sqrt{(x - \bar{x})^2}} \quad (10)$$

with  $y$  ( $\hat{y}$ ) observed (fitted) values of dependent variable,  $n$  number of observations,  $k$  number of variables and  $x$  the respective marketing variable.

### Insert Table 7 Here

Indeed, the marketing variables are statistically significant confirming that our model provides a better fit explaining the deviations from the curve (the purchase history) than a model with RFM only. The current effects' elasticities for TV and Internet advertising carry the expected sign. Based on a Meta study Tellis (2009) found that the average advertising elasticity is .1 but also states that this elasticity is lower in models that use disaggregate data and include advertising carryover, quality, or promotion. Considering the fact that we estimate

an individual-level model and incorporate carry-over effects into our advertising variables, we believe that these two elasticities of .030 and .069 lie within a plausible range. It also seems reasonable that Internet advertising has a higher elasticity given that we analyze purchase behavior of an online service, and the advertising takes place at the point of sale. The elasticity for radio advertising is negative but so small that it will not be part of further discussion. One reason for this non-plausible and close to zero elasticity might be the sparse and concentrated occurrence in the observation period (only four times) combined with the uncertainty of how many customers were actually exposed to a radio advertising campaign whereas it is unlikely to miss the broadly spread TV and Internet campaigns. An explanation for the negative, but also rather small coupon elasticity could be the fact that the ‘impulse’ (positive relative change) from coupons appears only in the first week of the campaign because it is dummy variable (for detailed construction of the impulse see 3.2, *model specification*). In case most customers react with a delay, which may be possible, the result would presumably be a negative elasticity. Finally, it must be taken into account that we investigate the elasticities with respect to purchase incidence instead of sales as is usually done; furthermore we analyze music downloads, which are innovative hedonic products instead of frequently purchased consumer goods, thus requiring a careful interpretation throughout.

The results of the negative binomial panel regression are presented in table 8.

### **Insert Table 8 Here**

The prediction from the incidence model has a significant, positive (and the strongest) effect on purchase quantity, confirming that the model in the first stage performs well in forecasting purchase probability. However, the monetary value shows a significant, negative coefficient in contrast to the incidence model. This implies that customers with low cumulative revenue tend to purchase more at a time than customers already being characterized by a high

monetary value. (Customers with high monetary value purchase more often with less quantity supporting preliminary analyses). With regard to marketing effects, TV advertising, again, has a significant and positive impact, and thus can be considered as an effective instrument for both purchase incidence and quantity. Internet advertising, here with a significant, negative effect, obviously helps to increase the purchase probability, but not to raise the quantity bought during one shopping trip. Radio and coupons are not significant, possibly due to the same data problems outlined above.

The customer-individual trend has a positive impact as opposed to the first stage, indicating that higher quantities are purchased with increasing relationship duration. A reason for this might be that customers get used to this download service over time and develop a certain loyalty, and are therefore tempted to buy more. The variable “coupon with first purchase” has the same negative sign as in the first stage. The remaining variables control for seasonality and holidays, and predominantly show significant effects similar to the ones in the incidence model.

#### *4.3 Validation and Forecasting Performance*

First, we present the results of the holdout validation in order to evaluate the out-of-sample forecast performance of our proposed modeling framework relative to the in-sample performance (e.g., Steckel & Vanhonacker, 1993) being important to assess the generalizability, stability, and robustness of the model while avoiding overfitting (Leeflang et al., 2000, p. 500-501). Contributing to the panel specification and with respect to managerial relevance, we perform a holdout validation over time, that is we use all weeks up to week 79 to calibrate the model (number of weeks varies for each customer) and the last eight weeks (up to week 87) to validate it. Based on the estimation results of the calibration period, we calculate the out-of-sample predictions and correlate them with the actual values.



Analogously, we take the in-sample predictions, i.e. the fitted values for week 80 to 87 based on the estimation of all 87 weeks, and correlate them with the actual values as well. The results for both model stages are shown in table 9.

**Insert Table 9 Here**

By comparing the two correlation coefficients of each model, we find very similar results for in- and out-of-sample-prediction confirming the suitability of the modeling framework for performing a real forecast within a horizon of two months relative to its benchmark, the in-sample performance (and under normal market conditions).

In order to evaluate the absolute forecast performance of the logit model, we also calculate the percentage of correctly and incorrectly classified observations in the holdout period. We again compare the out-of-sample to the in-sample prediction. As expected, the overall classification performance is significantly higher (76.1% versus 52.1%) for the in-sample forecast. The out-of-sample forecast overestimates purchase activity a lot more than the in-sample classification by wrongly assigning 45.7% of the “zero-observations” to the purchase-category, leading to a notably worse fit of the no-purchase category. However, the out-of-sample classification provides – with almost 70% – a considerably better fit in predicting the purchase event than the classification based on full information (58.2%). In fact, it is usually more important to predict actual purchases rather than non-purchases so that the worse overall result of the real forecast is outweighed by the better prediction of the relevant category.

**Insert Table 10 Here**

Finally, we test the performance of the quantity model based on the cumulative relative absolute error (CumRAE) with respect to a naïve forecast, defined as  $\hat{x}_t = x_{t-1}$  (Armstrong & Collopy, 1992). Due to its relative and standardized character providing a benchmark for

prediction quality (as opposed to the popular RMSE), this metric is particularly appropriate for our single model case (Barrot, 2009).

The test statistic is given by

$$CumRAE = \frac{\sum_{t=h+1}^T |\hat{x}_t - x_t|}{\sum_{t=h+1}^T |x_{t-h} - x_t|} \quad (8)$$

In case the result is approximately 1 the model prediction does not differ significantly from the naïve forecast, for values below (greater than) 1 the proposed model predicts better (worse). Applied to our quantity model, we obtain a value of .64 revealing a considerably better performance than the naïve forecast.

In summary, we can affirm a decent overall prediction performance especially when taking into account the strong heterogeneity of the customer database and the comparably simple estimation method by using a random effects model. Particularly the incidence model, being the focus of our study, demonstrates its robustness with respect to the validation period.

## 5. Discussion and conclusion

The focus of this study has been to develop a modeling framework for the prediction of repeat purchases in the online environment that adequately determines the impact of advertising and promotion, clearly differentiated from the impact of the purchase history, and allows for optimization of the marketing-mix. By postulating that RFM already captures the purchase process and marketing decision variables only shift this basic curve, i.e. RFM and marketing are highly correlated, we built a model that multiplicatively connects relative changes of marketing with purchase history variables (analogous to Bass, Krishnan, & Jain, 1994).

In fact, the significant results of the SUR model confirm the assumption that the purchase history already reflects customer response to a certain level of marketing activities. Consequently, marketing variables cannot be estimated independently from the purchase history because both variable categories include marketing information, making it impossible to separate the original marketing impact. A simple linear combination as applied in other studies (e.g., Ansari, Mela, & Neslin, 2008; Manchanda et al., 2006) would underestimate the effects of the marketing instruments leading to biased and/or non-plausible results. The hierarchical formulation (of two levels) suggested by Rust and Verhoef (2005) could cause the same biased and instable coefficients because the second level, i.e., the purchase history effects would partly offset the marketing effects at the first level in case the variables are significantly correlated across levels. Moreover, by multiplicatively connecting both variable categories as done by Khan, Lewis, and Singh (2009) it is possible to investigate how marketing effects change when they are linked to different levels of, e.g., recency and frequency. However, as long as absolute values are used for the marketing variables we can neither clearly separate the purchase history effects from the original marketing impact nor optimize the marketing-mix. The same problem would apply when using the modeling frameworks of Venkatesan and Kumar (2004) and Venkatesan, Kumar, and Bohling (2008) who include marketing and transactional information in separate sub-models multiplicatively linked together. Frameworks which incorporate the marketing instruments into a brand choice utility function, and model the purchase incidence solely as a function of the purchase history (e.g., Zhang & Krishnamurti, 2004; Zhang & Wedel, 2009) are also unsuitable for our objectives since hedonic media downloads such as music, books and movies, are unique. Given that a brand choice decision must not be made marketing should directly influence the purchase decision itself (with respect to both, incidence and quantity). In summary, neither of the relevant studies on repeat purchase behavior discussed above provides a modeling framework that matches all our conditions and assumptions like the model of Bass, Krishnan,

and Jain (1994) for an analogous problem does. Besides, providing an adequate basis is crucial also for optimization; if the model does not have the ability to estimate the original marketing effects, it is not possible to optimize correctly, leading, in turn, to a sub-optimal budget allocation. Our model formulation even allows for direct optimization without further transformation because the coefficients of the marketing variables already represent elasticities.

By using a—in our opinion—more adequate formulation we also expected to obtain a better model fit and forecast performance than achieved by the commonly employed model with the linear combination of RFM and marketing variables. In fact, the analysis reveals that our proposed model performs better than the model with the linear function, albeit only slightly. Moreover, the results show that the pure RFM-model performs almost as well as the model with marketing information which confirms previous findings on the forecasting performance of RFM; this also supports our hypothesis that past marketing influence is captured by purchase history variables.

The absolute prediction performance is somewhat moderate, which is not surprising due to the fact that we investigated digital hedonic products which are very different to consumer packaged goods, and are also not purchased on a regular basis as “necessary” CPGs are. Purchase behavior regarding music downloads is influenced by situational factors and moods rather than by rational motives and therefore must be considered as random to some extent. Consequently, it is quite challenging to accurately predict future purchases; this has to be kept in mind when evaluating the forecasting performance. By accounting for this strong uncertainty, we arrive at the conclusion that the modeling framework, i.e., the incidence model in particular, provides an acceptable overall prediction performance, thus helping managers to improve customer management processes and effectively implement marketing activities. Due to the hedonic nature of this market, wide-ranging advertising and promotion

campaigns (above average awareness) not only stimulate buying more, or at an earlier point in time as it is often the case in the CPG market, but can even create the need to make a purchase in the first place. Thus, it is substantial to adequately incorporate the marketing-mix into the purchase incidence model. On the other hand, purchase quantity is only partially affected by marketing activity; there, other factors such as the relationship duration and the level of past sales volumes are found to be more important. The significant difference between the effects of the incidence and the quantity model validates the separate analysis of both stages, which is consistent to previous studies. Interestingly, by showing a significant, positive effect in both stages TV, advertising still plays a major role for online repurchase behavior despite being a traditional mass media instrument. Admittedly, it is also the instrument with the highest budget allocation in this study which in turn leads to the greatest media coverage. Given that our modeling framework determines the original impact we can indeed confirm that TV advertising is effective for predicting purchases in this new online market of digital hedonic products. Thus, further investment into TV advertising is strongly recommended. Internet advertising only affects purchase incidence positively, but with a higher elasticity than TV. However, since it takes place in the same channel where media downloads are sold, such a finding is only reasonable. In fact, similar results regarding the effectiveness of TV versus Internet advertising have been found by Reimer, Rutz, and Pauwels (2010).

In conclusion, this study sheds light on important issues pertaining to the development of a repeat purchase model for digital hedonic products that adequately accounts for the impact of marketing. By proposing a particular modeling framework, we tended to accommodate the problems that occur when ignoring the explicit relationship between marketing and purchase history variables, and to provide a basis for optimizing the effectiveness of the marketing-mix at the same time.

The study has some limitations that can be addressed by future studies. First, in our analysis, we considered Internet only as a nominal variable due to missing information on GRPs. In addition, radio advertising was only placed sparsely; given that it is also a mass media instrument applied to individual purchase behavior, it remains unclear if the real effects were indeed captured. In order to cope with today's (customized) individual-level marketing, it would also be important to investigate price promotions and direct marketing instruments such as (customized) emails.

Moreover, the proposed repeat purchase model is applied to customers of a company in the music download industry, which is a new and innovative field. As one of the first studies in this field—to the authors' knowledge—and given the strong uncertainty inherent in predicting customer repurchase behavior for hedonic media downloads, the findings have to be interpreted cautiously. Further empirical analyses of other service providers selling digital hedonic products within the media download industry are necessary before the findings can be generalized, especially the elasticities, for the entire market. Also, future research studies need to investigate whether the modeling framework and the results can be applied to other industries and settings.

## 6. References

- Abe, M. (2009).** "Counting Your Customers" One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model. *Marketing Science*, 28 (3), 541-553.
- Allenby, G. M., Leone, R. P., & Jen, L. (1999).** A Dynamic Model of Purchase Timing with Application to Direct Marketing. *Journal of American Statistical Association*, 94 (June), 365–374.
- Andrews, R. L., & Currim, I. S. (2009).** Multi-stage purchase decision models: Accommodating response heterogeneity, common demand shocks, and endogeneity using disaggregate data. *International Journal of Research in Marketing*, 26 (3), 197-206.
- Ansari, A., Mela, C. F., & Neslin, S. A. (2008).** Customer Channel Migration. *Journal of Marketing Research*, 45 (1), 60-76.

- Armstrong, J. S., & F. Collopy (1992).** Error Measures for Generalizing about Forecasting Methods: Empirical Comparisons. *International Journal of Forecasting*, 8, 69-80.
- Batislam, E. P., Denizel, M., & Filiztekin, A. (2007).** Empirical validation and comparison of models for customer base analysis. *International Journal of Research in Marketing*, 24 (3), 201-209.
- Barrot, C. (2009).** Prognosegütemaße. In Albers, Sönke, Daniel Klapper, Udo Konradt, Achim Walter und Joachim Wolf (Hrsg), *Methodik der empirischen Forschung*. 3. überarbeitete und erweiterte Auflage, Wiesbaden 2009, 547-560.
- Bass, F. M., Krishnan, T. V., & Jain, D. C. (1994).** Why the Bass Model fits without Decision Variables. *Marketing Science*, 13 (3), 203-223.
- Borle, S., Singh, S. S., & Jain, D. C. (2008).** Customer Lifetime Value Measurement. *Management Science*, 54 (1), 100–112.
- Chintagunta, P. K., & Haldar, S. (1998).** Investigating Purchase Timing Behavior in Two Related Product Categories. *Journal of Marketing Research*, 35 (1), 43-53.
- Fader, P. S., & Hardie, B. G.S. (2005).** The Value of Simple Models in New Product Forecasting and Customer-Base Analysis. *Applied Stochastic Models in Business and Industry*, 21 (4-5), 461–473.
- Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005a).** Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science*, 24 (2), 275–284.
- Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005b).** RFM and CLV: Using Iso-Value Curves for Customer Base Analysis. *Journal of Marketing Research*, 42 (4), 414–430.
- Greene, W. (2003).** *Econometric Analysis*, Upper Saddle River, Prentice Hall.
- Gupta, S., Hanssens, D., Hardie, B. G. S., Kahn, W., Kumar, V. Lin, N., & Ravishanker, N. (2006).** Modeling Customer Lifetime Value. *Journal of Service Research*, 9 (2), 139-155.
- Gupta, S. & Zeithaml, V. (2006).** Customer Metrics and Their Impact on Financial Performance. *Marketing Science*, 25 (6), 718-739.
- Heckman, J. (1979).** The sample selection bias as a specification error. *Econometrica*, 47 (1), 153-62.
- IFPI Digital Music Report (2009).** New Business Models for a Changing Environment. *IFPI*, January 2009.
- Khan, R., Lewis, M., & Singh, V. (2009).** Dynamic Customer Management and the Value of One-to-One Marketing. *Marketing Science*, 28 (6), 1063-1079.
- Kinshuk, J., Fader, P. S., & Hardie, B. G. S. (2010).** New Perspectives on Customer 'Death' Using a Generalization of the Pareto/NBD Model. *Marketing Science*, forthcoming.

- Kumar, V. & Shah, D. (2009).** Expanding the Role of Marketing: From Customer Equity to Market Capitalization. *Journal of Marketing*, 73 (6), 119-136.
- Leeflang, P. S. H., Wittink, D. R., Wedel, M., & Naert, P. A. (2000).** *Building Models for Marketing Decisions*. Kluwer Academic Publishers, Dordrecht.
- Lewis, M. (2004).** The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention. *Journal of Marketing Research*, 41 (3), 281-292.
- Maddala, G. S. (1987).** Limited Dependent Variable Models Using Panel Data, *Journal of Human Resources*, 22 (3), 305-338.
- Manchanda, P., Dubé, J.-P., Goh, K. Y., & Chintagunta, P. K. (2006).** The Effect of Banner Advertising on Internet Purchasing. *Journal of Marketing Research*, 43 (1), 98-108.
- Naik, P. A., & Raman, K. (2003).** Understanding the Impact of Synergy in Multimedia Communications. *Journal of Marketing Research*, 34 (2), 248-61.
- Naik, P. A., Raman, K., & Srinivasan, S. (2007).** Modeling Corporate Advertising. Under review in *Journal of Marketing Research*, June 2007.
- Puhani, P. A. (2000).** The Heckman Correction for Sample Selection and Its Critique. A Short Survey. *Journal of Economic Surveys*, 14 (1), 53-68.
- Reimer, K., Rutz, O. J., & Pauwels, K. H. (2010).** Marketing Effectiveness in the Music Download Industry. *MSI working paper series*, [10-112].
- Reinartz, W. J., & Kumar, V. (2000).** On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing. *Journal of Marketing*, 64 (4), 17-35.
- Reinartz, W. J., & Kumar, V. (2003).** The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing*, 67 (January), 77-99.
- Reinartz, W. J., Thomas, J. S., & Kumar, V. (2005).** Balancing Acquisition and Retention Resources to Maximize Customer Profitability. *Journal of Marketing*, 69 (1), 63-79.
- Roel, R. (1988).** Direct Marketing's 50 Big Ideas, *Direct Marketing*, 50 (May), 45-52.
- Rust, R. T., & Verhoef, P. (2005).** Optimizing the Marketing Interventions Mix in Intermediate-Term CRM. *Marketing Science*, 24 (3), 477-489.
- Schmittlein, D. C., & Peterson, R. A. (1994).** Customer Base Analysis: An Industrial Purchase Process Application. *Marketing Science*, 13 (1), 41-67.
- Simon, H. (1982).** ADPULS An Advertising Model with Wearout and Pulsation. *Journal of Marketing Research*, 19, 352-363.
- Steckel, J. H., & Vanhonacker, W. R. (1993).** Cross-Validating Regression Models in Marketing Research. *Marketing Science*, 12 (4), 415-427.



- Tellis, G. J. (2009).** Generalizations about Advertising Effectiveness in Markets. *Journal of Advertising Research*, 49 (2), 240-245.
- Van den Poel, D., & Buckinx, W. (2005).** Predicting online-purchasing behavior. *European Journal of Operational Research*, 166 (2), 557-575.
- Venkatesan, R., & Kumar, V. (2004).** A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy. *Journal of Marketing*, 68 (October), 106-125.
- Venkatesan, R., Kumar, V., & Bohling, T. (2007).** Optimal CRM Using Bayesian Decision Theory. *Journal of Marketing Research*, 44 (November), 579-594.
- Wübben, M., & von Wangenheim, F. (2008).** Instant Customer Base Analysis: Managerial Heuristics Often "Get It Right". *Journal of Marketing*, 72 (3), 82-93.
- Zhang, J., & Krishnamurthi, L. (2004).** Customizing Promotions in Online Stores. *Marketing Science*, 23 (4), 561-578.
- Zhang, J., & Wedel, M. (2009).** The Effectiveness of Customized Promotions in Online and Offline Stores. *Journal of Marketing Research*, 46 (2), 190-206.

**Table 1**

Benchmark models for purchase incidence model.

	Specification of the logit utility function <sup>a</sup>
Model1	RFM + trend
Model2	Model1 + frequency <sup>2</sup>
Model3	Model2 + advertising-co + coupons
Proposed model	(RFM) * (1+ $\Delta$ advertising-co+ $\Delta$ coupons) + trend

<sup>a</sup> identical set of covariates will be added to each model**Table 2**

Descriptive statistics of purchase history variables.

per week	Mean	Median	Maximum	Minimum	SD
<i>Dep. var</i>					
Purchase incidence	.12	.00	1	0	.325
# Items	.55	0	122	0	2.48
<i>Predictors</i>					
Recency	11.31	7.00	80	1	11.44
Mean-adjusted recency	0	-1.32	49.20	-32.65	9.52
Frequency <sup>a</sup>	.28	.14	1	.01	.31
Frequency_square <sup>1</sup>	.18	.02	1	2e <sup>-4</sup>	.33
Monetary value	33.32	20.45	701.60	.01	42.71

<sup>a</sup> aggregation level is weeks, so maximum is 1.

**Table 3**

Descriptive statistics of marketing variables.

	Mean	Median	Maximum	Minimum	SD
Absolute values (GRPs for TV and radio) without carry-over (based on 87 weeks)					
TV	35.86	20	139	0	38.4
Radio	16.26	0	299	0	56.5
Internet	.53	1	1	0	.50
Coupons	.36	0	1	0	.48
Positive relative changes including carry-over (based on total sample, n=346,882)					
$\Delta$ TV-co	.03	0	1.34	0	.07
$\Delta$ Radio-co	33.60	0	2,213.22	0	268.95
$\Delta$ Internet-co	.09	.01	1.86	0	.20
$\Delta$ Coupons	.28	0	2	0	.70

**Table 4a**

Summary statistics of the SUR model.

Equation	RMSE	R <sup>2</sup>	Chi <sup>2</sup>
Recency (mean-adjusted)	8.529	.198	85621.98**
Frequency	.307	.045	16364.35**
Frequency <sup>2</sup>	.327	.026	9187.35**
Monetary value (mv)	41.111	.074	27580.69**

Breusch-Pagan test of independence: Chi<sup>2</sup>(6) = 4.78e+05, Pr = .0000

Note: \*\*p&lt;.01 significant

**Table 4b**

Estimation results of the SUR model.

<i>Dep. var</i>	Estimate (Std. Error)		<i>Dep. Var</i>	Estimate (Std. Error)	
Explanatory vars			Explanatory vars		
<i>Frequency</i>			<i>Frequency square</i>		
Constant	.468**	(.002)	Constant	.333**	(.002)
TV-co	-1.1E-04**	(3.3E-06)	TV-co	-1.0E-04**	(3.6E-06)
Radio-co	-3.5E-05**	(5.3E-06)	Radio-co	-3.5E-05**	(5.6E-06)
Internet-co	-.023**	(2.0E-04)	Internet-co	-.018**	(2.2E-04)
Coupons	.009**	(.001)	Coupons	.007**	(.001)
<i>Mean-centered recency</i>			<i>Monetary value</i>		
Constant	-8.087**	(.063)	Constant	3.132**	(.304)
TV-co	.001**	(9.3E-05)	TV-co	.016**	(4.5E-04)
Radio-co	.001**	(1.5E-04)	Radio-co	.005**	(7.0E-04)
Internet-co	1.603**	(.006)	Internet-co	4.115**	(.027)
Coupons	-.433**	(.032)	Coupons	-1.050**	(.153)

Note: \*\*p&lt;.01, ns not significant, two-tailed significance levels, N= 5,478

**Table 5**

Performance criteria for the different logit model specifications.

	Specification	LL	BIC	McFadden R <sup>2</sup>
<b>Model1</b>	RFM + trend + covariates	-116567.6	233,454	.083
<b>Model2</b>	Model1 + frequency <sup>2</sup>	-116013.6	232,359	.087
<b>Model3</b>	Model2 + advertising-co + coupons	-115936.5	232,256	.088
<b>Proposed model</b>	(RFM) * (1 + $\Delta$ advertising-co+ $\Delta$ coupons) + trend + covariates	-115751.1	232,089	.089

BIC =  $-2LL + K\ln(T)$ , where LL is the maximized log-likelihood value, T is the sample size, and K is the number of parameters.

**Table 6**

Estimation results of the proposed incidence model (full sample).

	Estimate	Std. Error
Constant	-2.855***	.037
<i>RFM</i>		
Recency (mean-centered)	.028***	.001
Frequency	4.947***	.126
Frequency <sup>2</sup>	-3.181***	.101
Monetary value (mv)	.002***	1.9E-04
<i>Marketing</i>		
$\alpha_0\beta_i$		
$\Delta$ TV-co	-.459**	.218
$\Delta$ Radio-co	3.3E-04***	4.6E-05
$\Delta$ Internet-co	.048***	.007
$\Delta$ Coupons	.054**	.022

Note: \*p<.1, \*\*p<.05, \*\*\*p<.01, ns not significant, two-tailed significance levels, N= 5,478

**Table 6**

Estimation results of the proposed incidence model (full sample) – continued.

	Estimate	Std. Error
<i>Interaction terms</i>	$\alpha_i \beta_j$	
Recency- $\Delta$ TV-co	-.007	.012
Recency- $\Delta$ Radio-co	5.0E-06*	2.8E-06
Recency- $\Delta$ Internet-co	-.001	.004
Recency- $\Delta$ Coupons	-.006***	.001
Frequency- $\Delta$ TV-co	-.822	1.315
Frequency- $\Delta$ Radio-co	4.7E-04*	2.5E-04
Frequency- $\Delta$ Internet-co	-1.902***	.384
Frequency- $\Delta$ Coupons	-.347***	.122
Frequency <sup>2</sup> - $\Delta$ TV-co	1.001	1.114
Frequency <sup>2</sup> - $\Delta$ Radio-co	2.2E-04	2.1E-04
Frequency <sup>2</sup> - $\Delta$ Internet-co	1.352***	.318
Frequency <sup>2</sup> - $\Delta$ Coupons	.237**	.102
MV- $\Delta$ TV-co	.014***	.003
MV- $\Delta$ Radio-co	4.0E-08	5.7E-07
MV- $\Delta$ Internet-co	.003***	.001
MV- $\Delta$ Coupons	2.2E-04	1.6E-04
<i>Covariates</i>		
Trend	-.016***	4.0E-04
Gender	.010	.017
Newsletter	-.017	.021
Permission	.011	.020
Coupon with 1 <sup>st</sup> purchase	-.199***	.035
Time from registration to 1 <sup>st</sup> purchase	.001***	.000
Album-release	-.011	.011
Single-release	.031***	.011
Holidays	-.029***	.010
February	.047	.034
March	.162***	.032
April	.164***	.035
May	.264***	.034
June	.305***	.030
July	.194***	.031
August	.115***	.030
September	.253***	.033
October	.314***	.033
November	.364***	.032
December	.222***	.032

Note: \*p&lt;.1, \*\*p&lt;.05, \*\*\*p&lt;.01, ns not significant, two-tailed significance levels, N= 5,478

**Table 7**

Inferred elasticities for marketing variables.

Marketing variable	Symbol	Current effects' elasticity (Std. Error)
$\Delta$ TV-co	$\beta_1$	.030*** (.008)
$\Delta$ Radio-co	$\beta_2$	-1.0E-04***(2.0E-06)
$\Delta$ Internet-co	$\beta_3$	.069*** (.003)
$\Delta$ Coupons	$\beta_4$	-.002** (.001)

Note: \*p&lt;.1, \*\*p&lt;.05, \*\*\*p&lt;.01, ns not significant, two-tailed significance levels, Ncust.= 5,478

**Table 8**

Estimation results of the quantity model (full sample).

	Estimate	Std. Error
Constant	-4.295***	.062
<i>Predictors</i>		
Prediction from logit model	8.071***	.091
Monetary value (mv)	-7E-04***	1E-04
TV-co	6E-04***	7E-05
Radio-co	1E-05	9E-04
Internet-co	-.030***	.005
Coupons	-.006	.013
<i>Covariates</i>		
Trend	.002***	6E-04
Album release	.013	.010
Single release	.006	.010
Gender	.010	.012
Coupon with 1 <sup>st</sup> purchase	-.120***	.026
Newsletter	-.036**	.015
Permission	-.011	.014
Time from registration to 1 <sup>st</sup> purchase	-.0002*	1E-04
Holidays	.017*	.010
February	.095**	.033
March	.063*	.036
April	.084**	.037
May	.120***	.036
June	.147***	.035
July	.257***	.040
August	.291***	.044
September	.134***	.037
October	.135***	.035
November	.036	.030
December	.070*	.038

Note: \*p&lt;.1, \*\*p&lt;.05, \*\*\*p&lt;.01, ns not significant, two-tailed significance levels, Ncust= 5,478

**Table 9**

Holdout validation.

Correlations for holdout period (weeks 80-87) <sup>a</sup>		
	Purchase incidence (logit)	Purchase quantity (NegBin given logit out of-sample prediction)
Out-of-sample predicted value <i>based on calibration sample</i>	.2371**	.1534**
In-sample predicted value <i>based on full sample</i>	.2538**	.1651**

<sup>a</sup>full sample (T=87 weeks) vs. calibration sample (T=79 weeks), predictions based on variables with t-value >=1; \*\*p<.01 significant

**Table 10**

Classification table of logit model for holdout period.

		Out-of-sample classification for weeks 80 - 87 (real forecast)			In-sample classification for weeks 80 - 87		
		Predicted values			Predicted values		
		<b>1</b>	<b>0</b>	Total correct per category	<b>1</b>	<b>0</b>	Total correct per category
Actual values	<b>1</b>	4.9% (2,165)	2.2% (959)	<b>69.3%</b>	4.1% (1,818)	3.0% (1,306)	58.2%
	<b>0</b>	45.7% (20,023)	47.2% (20,677)	50.8%	20.9% (9,183)	72.0% (31,537)	<b>77.5%</b>
Total correct		<u>52.1%</u>			<u>76.1%</u>		