Unveiling the Relationship between the Transaction Timing, Spending and Dropout Behavior of Customers

Nicolas Glady[1], Aurélie Lemmens[2] and Christophe Croux[34]

[1] Department of Marketing, ESSEC Business School

[2] Faculty of Economics and Business Administration, Tilburg

[3] Faculty of Business and Economics, K.U. Leuven

[4] This work is part of the Ph.D. thesis of the first author. All authors contributed equally to this paper. The project received financial support from the N.W.O. in the framework of a VIDI grant. All correspondence should be sent to Nicolas Glady: Department of Marketing, ESSEC Business School, Avenue Bernard Hirsch, BP 50105, 95021, Cergy-Pontoise, France; Glady@essec.edu, Phone: +33 1 34 43 28 95.

Unveiling the Relationship between the Transaction Timing, Spending and Dropout Behavior of Customers

Abstract

The customer lifetime value combines into one construct the transaction timing, spending and dropout processes that characterize the purchase behavior of customers. Recently, the potential relationship between these processes, either at the individual customer level (i.e. *intra-customer correlation*) or between customers (i.e. *inter-customer correlation*), has received more attention. In this paper, we propose to jointly unveil the direction and intensity of these correlations using copulas. We investigate the presence of these correlations in four distinct product categories, namely online music albums sales, securities transactions, and utilitarian and hedonic fast-moving consumer good retail sales.

For all product categories, we find a substantial amount of inter- and intra-customer correlation. At the inter-customer level, on average frequent buyers tend to spend more per transaction than the other customers. In addition, on average, large buyers have a longer lifetime. At the intra-customer level, we find that the existence and intensity of compensating purchase behaviors vary across product categories and across customers. From a managerial viewpoint, our approach improves the forecasts of the firm's future cash flows, especially for the product categories and customers where the correlations are the strongest. Moreover, the correlation parameters also provide additional insights to traditional customer valuation analysis on the magnitude, durability and volatility of the cash flows that each customer generates. We conclude by discussing how these insights can be used to improve customer portfolio decisions.

KEYWORDS: Correlation, Copula, Customer Lifetime Value, Customer Portfolio Management, Pareto/NBD Model, Volatility of Cash Flows.

1. Introduction

Over the last decade, customer lifetime value (hereafter, CLV) has become a powerful customer valuation metric (Blattberg et al. 2009, Gupta et al. 2004, 2006, Kumar and Reinartz 2006, Kumar et al. 2008, Rust et al. 2004, Venkatesan and Kumar 2004). Its success among academics and practitioners can be explained by the increasing pressure to make marketing accountable and the need to identify profitable customers and allocate resources accordingly (Gupta et al. 2006, Kumar and Reinartz 2006, Venkatesan et al. 2007). In this context, accurate CLV forecasting has become of upmost importance to managers.

The CLV metric integrates three key decision processes a customer goes through: (i) the transaction timing process, or *when to buy*, (ii) the spending process, or *how much to spend*,[1] and (iii) the dropout process, or *when to become permanently inactive* (Fader et al. 2005a). Together, these decisions determine the cash flows that firms can expect from each customer over her lifetime. These three purchase decisions have traditionally been assumed independent of each other (Fader et al. 2005a, Schmittlein and Peterson 1994). In particular, the model proposed by Fader et al. (2005a) assumes three underlying distributions that are independent: an exponential distribution for the customer's interpurchase times, a gamma distribution for the spend per transaction and an exponential distribution for the customer's unobserved lifetime. They assume customer heterogeneity in the various processes using independent mixing distributions. In practice however, various forms of correlation are likely to violate the independence assumption and consequently lead to inaccurate CLV forecasting. One type of correlation occurs at the *intra-customer* level when the timing at which a customer makes a

[1] The spending process captures how much a customer spends per transaction. Traditionally, the CLV framework considers the aggregate total (dollars) *value* of a transaction, i.e. the number of units bought per transaction (*quantity*) x the price per unit. For simplicity, we will refer to the notion of quantity per transaction or value per transaction interchangeably.

purchase interrelates with the value of this transaction. For instance, Jen et al. (2009) find evidence that some customers adjust their purchase quantities upwards when interpurchase times are longer. Another type of correlation occurs at the *inter-customer* level when the expected number of transactions, transaction value and/or customer lifetime correlate across customers. For instance, customers with a high purchase frequency have been found to generate greater income streams and have longer expected lives than those who purchase infrequently by Blattberg et al. (2001) and Jacoby and Kyner (1973).[2] In total, three inter-customer correlations – (i) between timing and spending, (ii) between timing and dropout and (iii) between spending and dropout – can arise between customers, as well as one intra-customer correlation for each individual customer.[3]

Lately, several methods have been proposed to unveil either the intra-customer correlation between the timing and spending processes (Boatwright et al. 2003, Glady et al. 2009, Jen et al. 2009, Romero et al. 2013), or the inter-customer correlation between the timing, spending and/or dropout decisions (Abe 2009a/b, Borle et al. 2008; we refer to the next section for a complete review). In all instances, research has shown that ignoring any of these correlations substantially biases the model predictions since it fails to account for the covariance between the processes when forecasting the CLV (Park and Fader 2004). However, to date no single study has been able to capture all correlation types at once.

In this paper, we propose a model for estimating CLV that jointly accounts both the intra- and inter-customer correlations between the timing, spending and dropout decisions of customers. Methodologically, we extend the model proposed by Fader et al. (2005a) by replacing

[2] Note that this concept of inter-customer correlation does not refer to any mechanism of influence or social contagion between customers.

[3] At the intra-customer level, the correlations between the dropout process and the other two processes (timing and spending) do not exist as permanent dropout only occurs once. Therefore, the total number of possible correlations is four, not six (three at the inter-customer level, one at the intra-customer level).

the independent distributions for each customer's interpurchase time and spend per transaction by a joint distribution (intra-customer level) and also specify a joint distribution for the transaction rates, spending rates and dropout rates across customers (inter-customer level). To link these distributions, we use *copulas* (Danaher and Hardie 2005, Danaher and Smith 2011). They are able to "couple" different families of distribution and to unmask the true strength of dependence between any two processes, which a classical correlation coefficient (e.g. Pearson) would not be able to identify. We show that accounting for both the intra- and inter-customer correlations improves predictions of customer purchase decisions, above and beyond incorporating none or some of them (intra or inter). Beyond the gains in predictive accuracy, we also discuss how the intra- and inter-customer correlations can guide customer portfolio decisions. Managerially, they contain useful information as to the magnitude, durability and volatility of the cash flows generated by every customer.

Finally, we also contribute to the customer valuation literature by developing a typology of the different correlations between the transaction timing, spending and dropout processes. We explain how they translate in terms of purchase behavior and why they are likely to occur. We focus on a number of rationales which underlie customer purchase decisions and create tradeoffs between the various decisions customers make (Chintagunta 1993, Gupta 1988). Finally, we ensure the generalizability of our findings by applying our model to four customer transaction databases, representing different product categories and/or industries. The first one pertains to music albums sales at an online retailer (CDNOW).[4] The second includes securities transactions at a major financial institution. The last two data sets concern the retail industry; one contains transactions of a utilitarian fast-moving consumer good (FMCG), the other of a hedonic FMCG.

[4] We thank Bruce Hardie who kindly provided us the data set.

The remainder of the paper is organized as follows. In Section 2, we define the intra- and inter-customer correlations between the timing, spending and dropout processes, review the existing literature, and discuss the behavioral rationales underlying each correlation. In Sections 3 and 4, we explain the copula methodology and show how to incorporate it into the CLV framework by Fader et al. (2005a). We describe our data in Section 5 and apply our methodology in Section 6. Sections 7 and 8 conclude with a number of managerial implications for customer portfolio management and a discussion of promising future research directions.

2. Inter- and Intra-Customer Correlation

In a non-contractual setting, we observe for each customer the timing and the value of her transactions (defined as the units bought per transaction x price per unit). Due to the non-contractual nature of the data, customers have an unobserved lifetime. The point at which a customer becomes inactive is estimated. Using this information, four correlations can be captured.

2.1. Definitions

Intra-customer correlation between the transaction timing and spending process. This correlation captures the degree to which the time preceding a given transaction (i.e. interpurchase time) relates to the value of this transaction. A positive correlation indicates that a customer who delays her purchase (i.e. shows a longer interpurchase time than usual) will buy in larger quantities than usual on that purchase. Respectively, a negative correlation indicates that a customer who delays her purchase will buy in smaller quantities than usual on that purchase. The intra-customer correlation is customer-specific.

Inter-customer correlation between the transaction timing and spending processes. This correlation captures the relation between the average number of purchases customers make (i.e.

individual transaction rate), and how much they spend on average per transaction (i.e. individual spending rate). A positive correlation indicates that customers who spend on average more per transaction than other customers also purchase more frequently than others. Respectively, a negative correlation indicates that customers who spend on average more per transaction than other customers purchase less frequently than others.

Inter-customer correlation between the transaction timing and dropout processes. This correlation captures the correlation between the average frequency at which a customer purchases and the hazard for this customer to become permanently inactive, given that the customer is still active (i.e. dropout rate). A positive correlation indicates that customers who purchase on average less frequently than others drop out later (that is, show a smaller dropout rate) than others. Respectively, a negative correlation indicates that customers who purchase on average more often than other customers drop out later than others.

Inter-customer correlation between the spending and dropout processes. This correlation captures the correlation between how much a customer spends on average per transaction and the hazard for this customer to become permanently inactive. A positive correlation indicates that customers who spend on average less per transaction than others drop out later than others. Respectively, a negative correlation indicates that customers who spend on average more per transaction than other customers drop out later than others.

2.2. Literature overview on the correlations

The correlations between the timing, spending and dropout decisions made by customers have received recent attention in the marketing literature. In Table 1, we provide an overview of the studies that have considered any of these correlations. For each study, we indicate the type of correlation considered, the dropout modeling context (i.e. contractual or non-contractual setting) and the industry under study.

[Insert Table 1 about Here]

At the intra-customer level, Boatwright et al. (2003) make purchase timing dependent on lagged quantity using a generalized Poisson specification. Jen et al. (2009) model the contemporaneous correlation between timing and quantity decisions using a bivariate log-normal distribution. In turn, Glady et al. (2009) link the expected number of transactions made by a customer and the expected value of these transactions by a dependency coefficient. At the inter-customer level, Fader et al. (2010) present an analysis that allows for inter-customer correlation between the transaction timing and dropout processes, and Abe (2009a) models this correlation using a bivariate normal distribution. Borle et al. (2008) and Abe (2009b) use a trivariate normal distribution to capture the correlation between all three processes. Finally, Schweidel and Knox (2013) use copulas to model the dependence between inter-donation times and donation amounts between and within customers, while Romero et al. (2013) use a partially hidden Markov model to forecast CLV while allowing the purchase timing and spending processes to be correlated between activity states at the intra and inter-customer level (keeping the independence assumption within a given state). Most of these studies are made in a non-contractual setting, except for Borle et al. (2008). While all these studies independently show that forecasts that take correlation into account outperforms forecasts than do not, none has consider all correlations simultaneously.

Finally, it is worth noting that, even though Fader et al. (2005a) assume the priors to be independent in their original model, it does not imply the posteriors to be independent as well (Fader et al. 2010).

2.3. Behavioral Interpretations and Underlying Rationales

Each correlation type and direction translates into a specific purchase behavior. In Table 2, we summarize the various cases and we discuss each of them in turn. For each, we also review a number of rationales to explain why and when these behaviors may occur.

[Insert Table 2 about Here]

2.3.1. Intra-customer correlation between the transaction timing and spending processes

A positive correlation reflects *compensating purchase behaviors* (Jen et al. 2009). Compensating behaviors indicate that a customer who hastens her purchase would compensate by buying in smaller quantities. In contrast, a negative correlation reflects temporary *purchase accentuation* (a customer who hastens her purchase would purchase larger quantities) *or attenuation* (a customer who defers her purchase would purchase smaller quantities).

The dependence between the timing and quantity decisions is commonly attributed to the presence of price promotions and other marketing activities (Gupta 1988). Customers often adjust their purchase timing and quantity decisions to the promotional calendar (Hendel and Nevo 2003, Van Heerde et al. 2004). For instance, customers sensitive to price promotions might delay their purchase (long interpurchase time) until the next promotion and buy more at the time of purchase (large quantities). Such behaviors can translate into compensating purchases.

In addition, the degree to which customers adjust their consumption to the level of their inventory (Ailawadi et al. 2007, Bell et al. 1999) can also affect the dependence between the timing and quantity decisions. In presence of inventory-based consumption (Ailawadi and Neslin 1998) or endogenous consumption (Sun 2005), customers tend to consume whatever quantity of products they purchased. Such behaviors might weaken the correlation between the purchase timing and quantity decisions.

Finally, the dependence between purchase timing and quantity decisions can be affected by changes in the marginal utility that a customer derives from a product. On the one hand, a positive state dependence caused by preference change (Dubé et al. 2010, Howard and Sheth 1969) and/or by customer learning (Gordon and Sun 2013) can lead to purchase accentuation. On the other hand, customers can experience wear-out in variety seeking (McAlister 1982) as their marginal utility to consume a given good decreases with consumption (Hartmann 2006), which can lead to purchase attenuation. These customers may satiate themselves with previous chosen goods or brands and switch brands in a quest for variety (McAlister 1982) or stop buying from the product category.

2.3.2. Inter-customer correlation between the transaction timing and spending processes

A positive correlation points to the existence of a *dual* customer base. A share of the customers buys more frequently and in larger quantities while another share buys at a lower frequency and in smaller quantities. In contrast, a negative correlation translates into *complementary* segments, in that a share of the customers buys in larger quantities but at a lower frequency than another share of customers who buys in smaller quantities but at a higher frequency. Finally, a non-significant correlation indicates that the amount a customer purchase on average is not informative as to how often she purchases and vice versa.

According to Baltas et al. (2010), the purchase timing and quantity decisions of customers depend on their acquisition and transaction utilities. Their household production framework explains how customers organize their shopping trips, and consequently their purchase behavior at the store (Ainslie and Rossi 1998, Manchanda et al. 1999), based on the benefits and costs of the products purchased (acquisition utility) and the (dis-)utility derived from the shopping activity itself (transaction utility; see Gupta and Kim 2010, Urbany et al.

1996, Vroegrijk et al. 2013, Chintagunta et al. 2012). In this framework, the relationship between the timing and spending processes depends on how the acquisition utility and transaction disutility correlate with each other. Product (categories) where the customers who derive the highest benefits from the product (high acquisition utility) are also those with extensive time and resources (low transaction costs) should show a positive correlation. Among other factors, differences in preference or attitudinal loyalty across customers can lead to such situations. Past research has shown that loyal customers tend to experience a higher acquisition utility and lower transaction costs than the less loyal customers (Krishnamurthi and Raj 1991, Krishnamurthi et al. 1992). In contrast, product categories where the customers that experience a high acquisition utility are also those who face high transaction costs should yield negative correlations. For instance, in the retailing sector, we often observe a segment of regular, *routine shoppers* (e.g. once a week trip) with large baskets, and a segment of frequent, *quick shoppers* (Bell and Lattin 1998, Kahn and Schmittlein 1989, Kim and Park 1997) with small baskets.

2.3.3. Inter-customer correlation between the transaction timing and dropout processes

A positive correlation points to the existence of *complementary* segments: lower-frequency longer-lived customers vs. higher-frequency shorter-lived customers. In contrast, a negative correlation translates into a *dual* customer base: lower-frequency shorter-lived customers vs. higher-frequency longer-lived customers. According to Reinartz and Kumar (2003), frequent customers tend to have longer lifetimes. As long as the interactions between the firm and the customers remain satisfactory, a high purchase frequency translates into a strong relationship between the firm and the customer, which leads to longer lifetimes (Morgan and Hunt 1994). Nevertheless, each transaction is also an opportunity for the customer to re-evaluate the utility derived from the product purchased. According to Fader et al. (2005b), a customer's probability to drop out is tied with her purchase occurrence. After any transaction, the customer "flips a coin" whether to become inactive with a certain probability. In this context, the more transactions, the higher the probability that the customer would drop out, suggesting a positive correlation. This effect is likely to be strengthened in a very competitive context where frequent buyers are more price-sensitive and demanding than the infrequent customers (Kim and Rossi 1994).

2.3.4. Inter-customer correlation between the transaction spending and dropout processes

Again, a positive correlation indicates *complementary* segments while a negative correlation translates into a *dual* customer base. The inter-customer correlation between the spending and dropout processes arises from differences in attitude towards the product (category) between customers, e.g. differences in their level of satisfaction, attitudinal loyalty or involvement. On the one hand, Jacoby and Kyner (1973), Bolton (1998) and Bolton and Lemon (1999) find that more satisfied customers have longer relationships with their service providers as well as higher usage levels of services. On the other hand, Reinartz and Kumar (2003) find customers who spend a large budget in a given product category to make more involved and elaborate purchase decisions than customers who spend small monetary amounts. According to Kim and Rossi (1994), involved customers are known to check competing offerings, are sensitive to their experience with the company and tend to re-evaluate their dropout decisions more substantively than small buyers.

In sum, at the inter-customer level, the direction and intensity of the relations between the timing, quantity and dropout decisions of customers result from differences between customers in their relationship with the product (category).

2.4. A Visual Comparison: Inter vs. Intra

In order to provide a visual representation of the difference between the inter- and intra-customer correlations between the timing and spending processes, we plot in Figure 1 the transaction value (in Euros) in function of the interpurchase time (in weeks) of transactions made by five imaginary customers. Transactions made by different customers are represented with different symbols, while the average interpurchase time and transaction value for each customer is depicted in bold.

[Insert Figure 1 about Here]

At the intra-customer level, three customers (the crosses, lozenges, and stars) show a positive correlation between their interpurchase times and transaction values, that is compensating purchases. In contrast, one customer (the squares) shows a non-significant correlation, while the remaining customer (the circles) shows a negative correlation, i.e. purchase accentuation or attenuation. At the inter-customer level, the bold symbols show a positive slope. Customers with longest interpurchase times, that is a lowest frequency (the diamonds and stars) offer the highest transaction values than customers with the shorter interpurchase times (the circles). As the average interpurchase time is inversely proportional to the frequency, the inter-customer correlation between the timing and spending processes is negative.

3. Copulas

As extensively described by Danaher and Smith (2011), a copula captures the association[5] between two random variables X and M with given marginal distributions $F(x) = P(X \leq x)$ and $G(m) = P(M \leq m)$. For example, F may be an exponential distribution and G a log-normal. The joint distribution function is given by $H(x, m) = P(X \leq x, M \leq m)$. Obtaining

[5] Strictly speaking, the word "association" is more appropriate than the word "correlation." The word "association" characterizes any type of relationship, whereas the word "correlation" applies strictly to linear relationships.

an explicit expression for the joint distribution H is generally cumbersome, motivating the use of copulas. The Sklar's theorem (Sklar 1959) yields that, for any F and G, there always exists a copula function C such that

$$H(x, m) = C(F(x), G(m)). \quad \textbf{(1)}$$

The copula function C is assumed to be known up to an unknown parameter θ. If f, g and h are the probability density functions corresponding to F, G and H, the copula density function c verifies

$$h(x, m) = c(F(x), G(m))f(x)g(m). \quad \textbf{(2)}$$

Various families of copulas exist. The simplest one is the independent copula, which assumes the independence between X and M, given by $C(F(x), G(m)) = F(x).G(m)$. The corresponding copula density is then equal to one. One can find a plethora of other copulas in the literature (see Nelsen 2006 for more detail). Among the most common, there is the Gaussian copula, which corresponds to a multivariate Normal distribution, and the Student copula corresponding to a multivariate t-distribution, allowing for heavier tails. In turn, the Gumbel copula only allows for a positive correlation between X and M, while the Frank copula does not allow for extreme correlation values. A visualization of these various copulas is provided in Figure 2, together with technical details in Appendix A. Figure 2 reports the contour plots corresponding to the joint distribution of two standard-normal random variables that have a Spearman rank correlation equals to .50, when specifying (i) an independent copula (upper-left plot), (ii) a Gaussian copula (upper-right plot), (iii) a Student copula (middle-left plot), (iv) a Gumbel copula (middle-right plot), and (v) a Frank copula (lower-left plot). While the joint distribution corresponding to the Gaussian copula is bivariate normal, the joint distribution corresponding to the Gumbel copula is asymmetric. In comparison to the Gaussian copula, the

Frank copula yields a weaker correlation in the tails, while the Student copula gives a heavier correlation in the tails. In the empirical application, we select the best copula based on its penalized goodness-of-fit. We use the Bayesian Information Criterion (BIC).

[Insert Figure 2 about Here]

One of the main advantages of copulas is their ability fit complex dependence structures without affecting the marginal distributions. This is particularly interesting when the distributions of the single random variables are well-known, as recently demonstrated in several marketing applications. For instance, Meade and Islam (2010) use copulas to model the dependence between successive interpurchase times. Meade and Islam (2003) and Sriram et al. (2010) introduce copulas to model the correlation between the times of adoption of related technologies. Another marketing issue where multivariate distributions prove useful is the modeling of household's purchase timing or incidence across related – complementary or co-incidental – product categories such as pasta and pasta sauce (Chintagunta and Haldar 1998), laundry detergent and fabric softener (Manchanda et al. 1999), or bacon and eggs (Danaher and Hardie 2005). Danaher and Hardie (2005) use the Sarmanov family of distribution (Sarmanov 1966), a special form of copulas, also used by Park and Fader (2004) and Danaher (2007) to model the correlation between multiple websites browsing patterns, and more recently by Schweidel et al. (2008) to account for the correlation between acquisition and retention times across customers. Danaher and Smith (2011) demonstrate that the Sarmanov family is more limited in its ability to model even moderate-sized correlation levels than the copulas described in this section. Recently, multivariate copulas have also been introduced in the marketing literature. Stephen and Galak (2012) model the correlation structure between traditional earned media, social earned media and sales while Kumar et al. (2014) use a Frank copula using a pair-copula construction method to model average transaction amounts jointly with opt-in and opt-out times to a firm's e-mail program. Finally, Park and Gupta (2012) recently proposed to use copulas to tackle endogeneity problems.

4. Customer Valuation using Copulas

We use copulas to model the intra- and inter-customer correlation between the timing, spending and dropout processes included in the customer valuation framework proposed by Fader et al. (2005a). They model the flow of transactions over time, accounting for dropout using a Pareto/NBD (Reinartz and Kumar 2000; 2003, Schmittlein et al. 1987, Schmittlein and Peterson 1994), as well as the monetary value of the transactions using a separate Gamma/Gamma sub-model.

For customer i, with $1 \leq i \leq n$, we observe transactions at times $t_{i,0} < t_{i,1} < \cdots < t_{i,x_i}$ with x_i the number of repeated transactions of customer i. The first transaction is made at time $t_{i,0}$ and the last transaction is made at $t_{i,x_i} \leq T_i$, where T_i indicates the time at which the CLV is computed. In the literature, t_{i,x_i} is referred to as the *recency* of customer *i*. The monetary values of the repeated transactions are given by $m_{i,1}, \ldots, m_{i,x_i}$. As done by prior work, the monetary value of the first transaction at $t_{i,0}$ is not used to model the CLV since it might be atypical of future purchases. For customer i, we compute the *interpurchase time* between transaction $j-1$ and j

$$IPT_{ij} = t_{i,j} - t_{i,j-1} \quad \text{for} \quad j = 1, \ldots, x_i, \tag{3}$$

and we denote R_i the time elapsed between t_{i,x_i} and T_i

$$R_i = T_i - t_{i,x_i}. \tag{4}$$

The assumptions on the marginal distributions of the spending process $m_{i,j}$, the interpurchase time process $IPT_{i,j}$, and the time to death are identical as in Fader et al. (2005a). The interpurchase time of a customer follows an exponential distribution with parameter λ_i, such that the total number of purchases in a unit time interval follows a Poisson[6] distribution with expected value λ_i. The dollar value of a customer's transaction follows a gamma (p, ν_i) distribution, with mean p/ν_i. Finally, the time to death (which is not observed) follows an exponential distribution with parameter μ_i, such that the expected lifetime equals $1/\mu_i$. We call λ_i the transaction rate, p/ν_i the spending rate, and μ_i the dropout rate of customer i.

Finally, we allow for both observed and unobserved customer heterogeneity as in Fader and Hardie (2007). We denote V_i^λ, V_i^ν, and V_i^μ the values of time-invariant covariates for customer i that we use to explain the transaction, spending and dropout rates. We specify

$$\lambda_i = \lambda_{0,i} \exp(\beta_\lambda V_i^\lambda) \quad \textbf{(5)}$$

$$p/\nu_i = (p/\nu_{0,i}) \exp(\beta_\nu V_i^\nu) \quad \textbf{(6)}$$

$$\mu_i = \mu_{0,i} \exp(\beta_\mu V_i^\mu), \quad \textbf{(7)}$$

where the random effect $\lambda_{0,i}$ follows a gamma(r, α) distribution, the random effect $\nu_{0,i}$ a gamma(q, γ) distribution, and the random effect $\mu_{0,i}$ a gamma(s, β) distribution. A positive β_λ indicates that the transaction rate increases when the value of a covariate increases. Similarly, a positive β_ν implies that the expected transaction value increases when the value of a covariate increases. Finally, a positive β_μ indicates an increase of the dropout rate when a covariate increases.

In the next sub-sections, we describe how we extend the original CLV model with copulas to model both the inter- and intra-customer correlations and allow for customer heterogeneity in the intra-customer correlation. A complete description of all model assumptions

---

[6] Other distributions, e.g. Weibull, log-logistic or log-normal, may be considered. For the purpose of comparison, we use the same specification as Fader et al. (2005a).

is given in Appendix B.

4.1. Modeling the Inter-Customer Correlation

At the inter-customer level, we capture the cross-sectional correlation between the transaction rate λ_i, the spending rate p/ν_i, and the dropout rate μ_i. Since the marginal distributions of the three rates were specified before, it remains to model the correlation using a trivariate copula c_{inter} with parameter $\theta_{inter} = (\sigma_{trans,drop}, \sigma_{trans,spend}, \sigma_{drop,spend})$, where $\sigma_{trans,drop}$ captures the correlation between the transaction rate λ_i and the dropout rate μ_i, $\sigma_{trans,spend}$ the correlation between the transaction rate λ_i and the spending rate p/ν_i, and $\sigma_{drop,spend}$ the correlation between the dropout rate μ_i and the spending rate p/ν_i. To do so, we extend equations (1) and (2) to the trivariate case. We model the trivariate correlation using elliptical copula, that is the Gaussian copula and the Student copula with 5 degrees of freedom, allowing for fatter tails. The Gumbel and Frank copulas result in expressions that are more complex to manipulate (Kumar et al. 2014). It is important to note that using a normal copula does not imply that the trivariate distribution is multivariate normal; the marginal distributions remain to be specified separately. The parameter of an elliptical copula is not a univariate scalar θ, but a matrix of the form

$$R(\theta) = \begin{pmatrix} 1 & \sigma_{trans,drop} & \sigma_{trans,spend} \\ \sigma_{trans,drop} & 1 & \sigma_{drop,spend} \\ \sigma_{trans,spend} & \sigma_{drop,spend} & 1 \end{pmatrix} \quad \textbf{(8)}$$

where the three parameters in (8) are between -1 and 1, but such that $R(\theta)$ is a positive definite matrix. In the implementation of the method, we reparametrized R as a function of three other parameters, all varying without restriction in the interval [0,1] and still ensuring positive definiteness of the correlation matrix (Pinheiro and Bates 1996).

4.2. Modeling the Intra-Customer Correlation with Heterogeneity

At the intra-customer level, we capture the correlation between $IPT_{i,j}$ and $m_{i,j}$ for each customer i through a copula c_{intra} with parameter $\theta_{intra,i}$. We parameterize the copula such that a positive $\theta_{intra,i}$ indicates that a transaction preceded by a longer interpurchase time (compared to the customer mean) is of higher value (compared to the customer mean). Similar to the transaction, spending and dropout rates, we allow for both observed and unobserved customer heterogeneity in the intra-customer correlation $\theta_{intra,i}$ by specifying $\theta_{intra,i}$ as a function of customer covariates V_i

$$f(\theta_{intra,i}) = \beta_{intra} V_i + \varepsilon_i, \qquad \textbf{(9)}$$

for a suitable link function f and with ε_i following a $N(0, \sigma^2_{intra})$. The link function ensures that θ_i stays within the bounds of the copula family.[7] The parameter β_{intra} captures the effect of a covariate on the strength of the intra-customer correlation.

4.3. Estimation of the Model Parameters

The parameters to be estimated (listed in Appendix B) are collected in a vector $\boldsymbol{\theta}$. The log-likelihood of $\boldsymbol{\theta}$ given $IPT_{i,j}$, T_i and $m_{i,j}$, for $1 \leq i \leq n$, and $1 \leq j \leq x_i$, is given by

$$\log L = \sum_{i=1}^{n} \log \int L_i(\lambda, \mu, \nu, \theta_{intra}) g_i(\lambda, \mu, \nu, \theta_{intra}) d\lambda \, d\mu \, d\nu \, d\theta_{intra}, \qquad \textbf{(10)}$$

where g_i is the joint distribution of the customers' individual parameters, and $L_i(\lambda,\mu,\nu,\theta_{intra})$ is given by

$$L_i(\lambda, \mu, \nu, \theta_{intra}) = \prod_{j=1}^{x_i} f(IPT_{ij}, T_i | \lambda, \mu) g(m_{ij} | \nu) \, c_{\theta_{intra}}\big(F(IPT_{ij}, T_i | \lambda, \mu), G(m_{ij} | \nu)\big).$$

Here, $f(\cdot \,|\lambda, \mu)$ and $F(\cdot \,|\lambda, \mu)$ are the density and c.d.f. of the distribution of the interpurchase

---

[7] In particular, $f(\dots) = 2/\pi \arctan(\dots)$ for the Gaussian copula, $f(\dots) = \exp(\dots) + 1$ for the Gumbel copula and $f(\dots) = \cdots$ for the Frank copula.

times and the time since the last transaction, and g and G are the density and c.d.f. of the distribution of the monetary values. The latter expression can be factorized as a product of three terms:

$$L_i(\lambda, \mu, \nu, \theta_{intra}) = A_i(\mu, \lambda) B_i(\nu) C_i(\lambda, \mu, \nu, \theta_{intra}).$$

For the first term $A_i(\mu, \lambda)$, we use (A1) and (A2) from Appendix B, and the properties of the exponential distribution to obtain (see Section 3.1 of Fader and Hardie 2005 for the derivation)

$$A_i(\mu, \lambda) = \frac{\lambda^{x_i+1}}{\mu+\lambda} \exp(-(\lambda+\mu)T_i) + \frac{\mu\lambda^{x_i}}{\mu+\lambda} \exp\big(-(\lambda+\mu)t_{x_i}\big). \quad \textbf{(11)}$$

For the second term, $B_i(\nu)$, we use (A3) and denote $\bar{m}_i$ the average of the monetary values of the repeated transactions and $\overline{\log(m_i)}$ the average of their log-transformations, such that

$$B_i(\nu) := \prod_{j=1}^{x_i} g(m_{ij}|\nu)$$

$$= \prod_{j=1}^{x_i} \frac{m_{ij}^{p-1} \exp\big(-m_{ij}\nu\big)\, \nu^p}{\Gamma(p)}$$

$$= \frac{\prod_{j=1}^{x_i} m_{ij}^{p-1} \exp(-x_i \bar{m}_i \nu)\, \nu^{p x_i}}{\Gamma(p)^{x_i}}.$$

Hence,

$$\log B_i(\nu) = (p-1)\overline{\log(m_i)} - x_i \bar{m}_i \nu + p x_i \log(\nu) - x_i \log\Gamma(p). \quad \textbf{(12)}$$

Finally, the third term $C_i(\lambda, \mu, \nu, \theta_{intra})$ is given by

$$C_i(\lambda, \mu, \nu, \theta_{intra}) = \prod_{j=1}^{x_i} c_{\theta_{intra}}\big(F(IPT_{ij}, T_i|\lambda, \mu), G(m_{ij}|\nu)\big). \quad \textbf{(13)}$$

Following the semi-parametric maximum likelihood approach for copula estimation (Genest et al. 1995), we approximate C_i by

$$\hat{C}_i = \prod_{j=1}^{x_i} c_{\theta_{intra}}(\hat{F}_{ij}, \hat{G}_{ij}),$$

where $\hat{F}_{ij}$ is the rank of IPT_{ij} among all other interpurchase times of the same customer, and similarly for $\hat{G}_{ij}$. Since the above quantity only depends on θ_{intra}, and A_i and B_i only depend on the other individual parameters, we can decompose the log-likelihood in (10) as follows:

$$\log L \approx \sum_{i=1}^{n} \log \int A_i(\lambda,\mu)B_i(\nu)g_i(\lambda,\mu,\nu)d\lambda\, d\mu\, d\nu \; + \log \int \hat{C}_i(\theta_{intra})g_i(\theta_{intra})d\theta_{intra}.$$

Since the expression above involves an integration, which cannot be solved analytically, we use Simulated Maximum Likelihood (SML) (see e.g. Green 2003, pp. 590-594). We draw $(\lambda_i^s, \mu_i^s, \nu_i^s, \theta_{intra,i}^s)$, for $s = 1, \ldots, S = 1000$ from the joint distribution of the individual parameters, and approximate

$$\log L \approx \sum_{i=1}^{n} \log(\frac{1}{S}\sum_{s=1}^{S} A_{is}B_{is}) + \sum_{i=1}^{n} \log(\frac{1}{S}\sum_{s=1}^{S} \hat{C}_{is}), \tag{14}$$

where A_{is}, B_{is}, C_{is} are defined in (11), (12), and (13), taking $\lambda = \lambda_i^s, \mu = \mu_i^s, \nu = \nu_i^s$ and $\theta_{intra} = \theta_{intra,i}^s$. The parameters involved in the first term of (14) are different from those in the second term, allowing for two separate maximization problems. The fact that the likelihood can be split in two parts allows us to study both levels of correlation separately.

We estimate $\boldsymbol{\theta}$ by maximizing the simulated log-likelihood in (14). Standard errors are retrieved from the Hessian of the log-likelihood. The random generation of the $(\lambda_i^s, \mu_i^s, \nu_i^s, \theta_{intra,i}^s)$ is based on the same random uniform number for every s, ensuring enough smoothness of the log-likelihood function. Once the (hyper-) parameters in $\boldsymbol{\theta}$ are estimated, following the empirical Bayes principle, we simulate from the posterior distribution of the individual parameters using the independent Metropolis-Hastings algorithm (see Appendix C for more detail). We obtain a point estimate for $\hat{\theta}_{intra,i}$ as the average of the outcomes of the simulated Markov Chain. Note that the estimation of the intra-customer correlation requires individual transaction data. When individual transaction data are not available, only the inter-customer correlation can be estimated. Moreover, the model boils down to the traditional Pareto/NBD model that for customers making no repeat purchase.

4.4. Forecasting

The CLV at horizon H for customer i is given by the discounted sum of all profits generated during the next H time units by this customer. Note that, in order to measure the entirety of the customer future activity, the prediction horizon H should be infinite. In this paper, we use an arbitrary limited horizon as in Reinartz and Kumar (2000) and Thomas (2001) in order to compare our predictions with the actual values. Strictly speaking, our measure is a "truncated CLV" which captures the discounted conditional expected total margin over a fixed time horizon. Let T denote the time at which the prediction is made. Future transactions take place at times $T + t_1, T + t_2, ...$ with corresponding transaction values $m_{\mathrm{i},t_1}, m_{\mathrm{i},t_2}, ...$ Adapting Gupta et al. (2004), we define the truncated CLV at horizon H as

$$TCLV_{i,H} = \sum_{t_j \leq H} \frac{m_{\mathrm{i},t_j}}{(1+d)^{t_j}}, \qquad \textbf{(15)}$$

where d is the discount rate. Note that the $TCLV_{i,H}$ defined in (15) is a random variable. Once the model is estimated, it is possible to simulate future transactions for every customer, resulting in the simulated distribution of the truncated CLV over the next H periods. Details are provided in Appendix C. The average over the simulated distribution yields a prediction of the expected truncated CLV for a given customer.

5. Data

Below, we describe each of the four data sets. Summary statistics are provided in Table 3.

[Insert Table 3 about Here]

5.1. Online Music Album Sales

The CDNOW data contains the individual transactions of 2,357 customers on the online music site CDNOW. The data cover 78 weeks for a sample of customers who made their first-ever purchase at the website during the first quarter of 1997. We use the first 39 weeks of 1997 for the estimation of the model parameters and keep the remaining 39 weeks as hold-out sample to assess the predictive performance of the model.

5.2. Financial Securities Transactions

The securities transaction data are provided by a major international financial service institution.[8] The data contain securities transactions made by 3,472 randomly-selected customers who made their first transaction between January 2001 and December 2003. Transactions include the purchase and selling of stocks at this bank between January 2001 and December 2005.[9] Based on a discussion with the firm management, we consider a profit margin of 0.1% of the transaction and an annual discount rate of 12%. We keep the last two years of transactions (from January 2004 to December 2005, $H = 24$ months) as hold-out sample to assess the predictive performance of the models. The data also contain socio-demographics and company-related customer characteristics that we use to model observed customer heterogeneity. Customer characteristics include a continuous customer age variable (mean-centered in the

[8] Due to data confidentiality agreements, we are unable to divulge more details about the companies providing us the data on the financial securities, utilitarian FMCG and hedonic FMCG.
[9] Note that we remove the stock transactions part of an automated pension plan from the data set.

model estimation), as well as a dummy variable *living area* accounting for the type of area where a customer is living. This variable takes the value one when the customer lives in the suburb of a city, and zero otherwise. As company-related customer covariates, we include a customer current lifetime variable (also mean-centered in the model estimation) that captures the number of weeks that the customer is with the firm (measured as the time since her first transaction). In addition, we also include a dummy taking value one when the bank is the customer's primary bank.

5.3. Utilitarian Fast-Moving Consumer Good

The third data set is provided by a major packaged goods manufacturer and contains transactions of a FMCG that can be classified as utilitarian product (e.g. detergent, toilet paper, pet food) according to Dhar and Wertenbroch (2000) or Khan et al. (2005). The data contain purchases made between March 2007 and December 2009 by the 2,968 customers who reported their first transaction to the data provider between March 2007 and December 2008. All these customers live in the same European country. We keep the last year of purchase (from January 2009 to December 2009, $H = 12$ months) as hold-out sample. The data also contain the size of the household (mean-centered in the model estimation) that we use to model observed customer heterogeneity.

5.4. Hedonic Fast-Moving Consumer Good

The fourth data set is provided by the same packaged goods manufacturer than the prior FMCG data set and contains transactions of a FMCG that can be classified as a hedonic product (e.g. ice cream, chocolate bar, wine, beer) according to Dhar and Wertenbroch (2000) or Khan et al. (2005). The data contain purchases made between January 2007 and August 2011 by the 5,682 customers who made their first transaction between January 2007 and December 2008. These customers live in the same European country as for the other FMCG. We keep the last 20 months of purchase (from January 2010 to August 2011, $H = 20$ months) as hold-out sample. The data also contain the size of the household (mean-centered in the model estimation).

6. Empirical Findings

For each data set, we estimate a model without and a model with inter- and/or intra-customer correlation. First, we assess whether the various data contexts show evidence of significant intra- and/or inter-customer correlations by model fit comparison. Second, we report the parameter estimates and the forecasts obtained by the best-fitting models.

6.1. Model fit comparisons

Model fit comparisons are reported in Table 4 (in bold, the best-fitting model). For each specification, we select the copula (Gauss, Student, Gumbel or Frank) that provides the best in-sample penalized fit using the BIC.

[Insert Table 4 about Here]

For all the datasets, we find that both levels of correlation improve the model’s penalized fit and are jointly significant (according to a likelihood ratio test comparing a model with both correlations vs. the null model without correlation). The following copulas at the intra- and inter-customer levels offer the best fit: Frank & Gauss for the online music album sales, Gauss & Student for the financial securities, Student & Gauss for the utilitarian FMCG, and Frank & Gauss for the hedonic FMCG. The Gumbel copula never performs best, indicating that the distribution of the correlation is relatively symmetric for all data sets. As a robustness check, we also estimated all models without covariates and found that the models with covariates and both the intra- and inter-customer correlations yield the lowest BIC in all cases. In Appendix D, we report all parameter estimates with respective standard errors, obtained for the best-fitting models. Below, we discuss the specific correlation results.

6.2. Estimated intra-customer correlations between timing and spending

In Figure 3, we plot the distribution of the intra-customer correlation for all the data sets. The x-axis represents the spearman correlation of the intra-customer correlation.

[Insert Figure 3 about here]

For the online music sales at CDNOW (Figure 3a), the intra-customer correlation is significantly positive (at the 10% probability level) but small for all the customers. There is little heterogeneity between customers as the small $\hat{\sigma}_{intra}$ = 0.14 (standard deviation = .26) in Appendix D testifies. Our findings are in line with those of Boatwright et al. (2003) for an online grocery retailer. The compensating purchase behavior of the customers in this product category is potentially driven by the fact that customers adjust their purchase timing and quantity decisions to the calendar of price promotions.

For the financial securities (Figure 3b), we find both positive and negative correlations between the interpurchase time and the amount spent over time. There is a substantial amount of unobserved heterogeneity in the intra-customer correlation as the large $\hat{\sigma}_{intra}$ = 1.85 (standard deviation = .07) testifies. We find that 20% of the population exhibits compensating purchase behavior (i.e. a significantly positive correlation at the 10% probability level), and 12% shows purchase accentuation or attenuation (i.e. significantly negative correlation). The remaining 68% have a non-significant correlation. The age of the customer has a significant effect (= -.38, standard deviation = .12) on this correlation. Young customers are more likely to show compensating purchase patterns, possibly because of the higher opportunity costs (e.g. time or budget constraints) they face. In contrast, older customers tend to show purchase accentuation, possibly because they are gradually learning how to make good investment decisions (learning effects); or purchase attenuation, either because their budget or interest (e.g. wear-out or satiation effect) in the stock market decreases over time.

For the utilitarian product (Figure 3c), we find a significant positive (at the 10% probability level) correlation for 15% of the customers and a significant negative correlation for only 5% of the customers. The individual consumption of this category is by nature non-expandable (e.g. toilet paper consumption) and not subject to inventory-based consumption. The product we investigate has a long expiration time (1 to 5 years in this product category) and its demand is known to be sensitive to price promotions. This explains compensating behaviors. Interestingly, household size has a positive effect on the correlation. Larger households might have larger storage capacity, and might therefore be more inclined to compensating behavior. A fraction of the customers show purchase accentuation or attenuation. They can be customers who enter or quit the product category (from or to a substitute category[10]). Interestingly, the correlation is similar to Jen et al. (2009) in another utilitarian category (office supply products).

For the hedonic good (Figure 3d), only a small fraction (about 6%) of the customers show a significant correlation at the 10% probability level. Hedonic products are subject to inventory-based consumption (Ailawadi and Neslin 1998). They are harder to resist and their purchase and consumption is more impulsive than the consumption of utilitarian goods. When bought in larger quantities, customers consume more of it. As a consequence, the quantity bought has little influence on the next purchase timing.

6.3. Estimated inter-customer correlations between timing and spending

In Table 5, we report the parameter estimates that capture the inter-customer correlation between the transaction rate (λ_i), spending rate (p/ν_i) and dropout rate (μ_i), as well as their standard errors (equation (3)). Overall, the four product categories offer consistent results.

[10] Due to data confidentiality, we cannot reveal the product category but it is useful to mention the existence of an important substitute category.

Customers show a positive inter-customer correlation between their timing and spending decisions, a positive or non-significant correlation between the timing and dropout decisions, and a negative or non-significant correlation between the spending and dropout decisions.

For all four product categories, higher purchase frequencies are associated with larger purchase amounts. The correlations are the strongest for both FMCGs. Although these results might first seem counterintuitive (Bell and Lattin 1998, Kahn and Schmittlein 1989, Kim and Park 1997), they are in line with the positive correlation found by Abe (2009b) for a large chain of music CDs. They contrast with the negative correlation Abe found for a Japanese department store in the same study. As pointed out by Abe (2009b), the correlations are context-dependent as they depend on how the acquisition utility and transaction disutility correlate with each other. Our results could possibly be explained by the fact that for the product categories involved, the customers include a large group of loyal and highly-involved customers and/or a large group of very sporadic, low-involved customers (Vilcassim and Jain 1991).

[Insert Table 5 about Here]

6.4. Estimated inter-customer correlations with dropout

The correlation between the customers' propensity to drop out and their other purchase decisions, i.e. when to buy and how much to buy, is broadly consistent across the four product categories. Overall, we find that frequent buyers have, on average, a higher propensity to dropout (a shorter lifetime) than the other customers. The correlation is significant for the music albums sales and for the financial securities, but not significant for both FMCG categories. The positive correlation supports the idea that each purchase is an opportunity for the customer to re-evaluate her choice for a certain brand or store (Fader et al. 2005a). The non-significant correlation for the FMCGs confirms the results of Abe (2009a, b). The correlation with spending is negative and significant for three of the product categories. The negative correlation corroborates the findings of Bolton and Lemon (1999) and Jacoby and Kyner (1973) that large buyers tend to be more involved and loyal in their purchase decisions than small buyers, as the budget they allocate to these purchases is larger.

6.5. Forecasting Results

We evaluate the predictive performance of the model that accounts for the intra- and inter-customer correlations vs. the model that assumes independence between the timing, spending and dropout decisions. For all four data sets, we make the predictions on a separate hold-out sample (see data section). We calculate the accuracy of both models in forecasting the truncated CLV over a fixed time horizon, as defined in the data section, using the root mean squared error (RMSE).

Boatwright et al. (2003) explain that a model with correlation would more accurately represent the customers in the tails of the frequency and spending distributions. Hence, our model should offer better forecasts for the customers with the highest actual CLV. Moreover, it should also perform best on the customers with the highest intra-customer correlation. Therefore, we compute the RMSE for: (i) all the customers, (ii) the top 10% customers with the highest estimated intra-customer correlation, (iii) the top 10% customers with the highest actual truncated CLV, and finally (iv) the customers selected in both subsamples (ii) and (iii) (that is, the customers who have a high intra-customer correlation and who are also in the tails of the frequency or spending distributions). In Table 6, we report, for all product categories, the RMSE for both models and the four different subsamples (i to iv).

[Insert Table 6 about Here]

Modeling correlation improves forecasts in most of the cases. As expected, the difference is more pronounced for the customers who show the strongest intra-customer correlation (column ii) and/or who are in the tails of the frequency or spending distributions (column iii). For the customers with an intra-customer correlation in the top 10% of the customer base and who are also located in the tails of the frequency or spending distributions (column iv), the copula approach shows the best predictive performance. While the results are good for the FMCGs, they remain mitigated for the financial securities and for the CDNOW data, which can be explained by the relatively low correlation in these data sets. However, even though the gains in predictive performance are not large, the benefits of modeling the correlation also lies in the understanding of the behavioral rationales underlying the correlations as well as in the additional insights that the correlation offers. The latter are described next.

7. Managerial Recommendations

The traditional approach to customer lifetime valuation focuses on maximizing the long-term cash flows of the firm. By forecasting the future stream of cash flows generated by customers, existing studies are able to rank and select customers based on the estimated net present value of their future cash flows, or alternatively based on the estimated change in their CLV after a marketing intervention. The most attractive customers are then targeted with specific marketing actions (Reinartz and Kumar 2000, 2003, Reinartz et al. 2005, Venkatesan and Kumar 2004). In this context, the main managerial value of CLV models investigated so far has been their ability to accurately predict CLV (see e.g. Fader et al. 2005a).

Beyond improving the accuracy of CLV predictions, unveiling the intra- and inter-customer correlation in CLV offers two additional metrics for valuing customers. In particular, the intra-customer correlation provides information on the *volatility* of the cash flows generated by a customer, while the inter-customer correlation sheds light on the tradeoff between the *durability vs. magnitude* of these cash flows. In this section, we discuss how these metrics enrich the traditional view on customer valuation. To illustrate the discussion, we focus on the utilitarian product category, which shows significant intra-customer and inter-customer correlations.

Cash flow volatility. Customers with stable cash flows are easier to manage than customers with highly volatile cash flows (Dhar and Glazer 2003, Fischer et al. 2013). Cash flow volatility impacts the firm's cost of capital (Rao and Bharadwaj 2008, Hanssens et al. 2009). It complicates production planning, increases inventory costs and leads to liquidity issues (Srinivasan and Hanssens 2009). Therefore, a proper customer valuation analysis does not only consider which customers generate the highest CLV but also takes into account the volatility of their cash flows (Gupta 2009). Citing Dhar and Glazer (2003), big spenders are – everything else equal – better than small ones, but what if the big spenders have extremely volatile spending streams while the small ones offer more stable cash flows?

The intra-customer correlation provides information on the volatility of customers' cash flows. By nature, compensating purchase patterns lead to more stable income streams than purchase accentuation or attenuation patterns. Customers with a positive intra-customer correlation compensate for late orders by ordering larger quantities (Jen et al. 2009). Customers with a negative correlation generate more variable income streams. When they buy earlier, they buy more; in the same way, when they buy later, they buy less. For the utilitarian FMCG, we find that customers with a significant negative correlation show on average 20% more volatile cash flows (average standard deviation = 829, average coefficient of variation = 1.10) than customers with a significant positive correlation (average standard deviation = 690, average coefficient of variation = .76). Firms can decide to favor steady cash flows by selecting customers with a positive intra-customer correlation. Alternatively, they can combine customers, or customer segments, with a negative intra-customer correlation if they have volatile but offsetting cash flow patterns (Tarasi et al. 2011, 2013).

Cash flow durability vs. magnitude. The durability of the cash flows generated by a customer is another important metric when valuing customers. Companies are urged to consider the tradeoff between the short-term and long-term sustainability of their marketing strategies (Hanssens and Dekimpe 2012). A portfolio of customers should ensure sufficient cash flows in the short and in the long run, by possibly combining customers who generate high cash flows and customers who have a long expected lifetime. The inter-customer correlation provides information on the tradeoff between the magnitude of the revenue stream generated by customers (timing x spending) and the expected durability of this stream (dropout).

When dropout decisions are negatively associated with the timing/spending decisions, as for the utilitarian FMCG, the largest buyers are also the ones with the longest expected lifetime. They contribute to the short-run and long-run sustainability of the firm's cash flows. In Table 7, we divide the population of repeat customers in the utilitarian product category (2466 customers) based on their estimated frequency and spending in four groups according to the median levels of both variables (i.e. a frequency of one transaction every 1.90 month and an average transaction value of 7.10 euros). Given the positive correlation, the large majority of customers (1902 customers) are classified as infrequent and small buyers (i.e. less than one transaction every 1.9 months and less than 7.10 Euros per transaction) or frequent and large buyers (more than one transaction every 1.9 months and more than 7.10 Euros per transaction). The number of customers in the other groups is comparatively very small.

[Insert Table 7 about Here]

For each group, we calculate the durability and magnitude of the cash flows. We use the median estimated lifetime as a measure of durability. For the magnitude, we compute the average estimated cash flows generated by the customers while being active as the average of the posterior means of the ratio $p/(\nu_i\lambda_i)$. Frequent and large buyers have, on average, longer lifetimes with the firm than the infrequent and small buyers (29.67 months compared to 18.45 months). Moreover, their cash flows are also of higher magnitude (i.e. 21.55 Euros per month) than those of the infrequent and small buyers (i.e. 13.05 Euros per month). Interestingly, we find that 22% of the frequent and large buyers have a positive intra-customer correlation. Focusing on this specific group should allow firms to maximize the magnitude and durability of its revenues while minimizing the volatility of its cash flows.

Managing a sustainable customer portfolio is more difficult when dropout decisions are positively associated with the timing/spending decisions. Customers who bring largest revenues in the short run (i.e. high-frequency, high-spending) are then also the ones with the shortest lifetime. In this case, firms have to sacrifice short-run incomes if they want ensure a more durable influx of cash flows. A solution can be to combine long-lived customers who generate smaller cash flows with a more frequently renewed set of shorter-lived customers who generate higher cash flows.

8. Conclusions, limitations and future research

While the CLV framework developed by Fader et al. (2005a) has been adopted as a powerful customer valuation method, the potential correlation between the timing, the spending and the dropout processes, both at the intra- and inter-customer levels, called for the development of a conceptual and methodological framework to better understand the trade-offs and interplay between the various purchase decisions that customers make. Using copulas to account for correlation, we offer a method to improve the model predictions and generate new insights into the purchase behavior of customers.

Despite our efforts, we should mention some limitations. First, while we explore the correlation across a variety of product categories, the set of variables available in each application is limited. We do not have access to marketing-mix variables, neither to competition variables, which would be interesting to consider in a future analysis. For the financial securities data, we do not have stock market fluctuation data, which may admittedly have a sizeable influence on customers' purchase and selling behavior. For instance, when the stock market is on the upward trend, customers might make more frequent and larger transactions. This omitted variable may drive both the timing and spending processes and may therefore inflate the intra-customer correlation. Addressing this problem is not trivial given that stock market fluctuations are difficult to predict. Second, our method captures the contemporary correlation between the various purchase decisions but overlooks the possibility of lead-lag relationships between the variables. Future research could benefit from adapting our framework to a context where longer lags would be added. Third, the range of copulas could be extended. For instance, pair-copula construction method could have been used (Kumar et al. 2014). Finally, our managerial recommendations based on the magnitude, durability and volatility of the cash flows generated by each customer constitutes a first step towards a different approach to customer portfolio decisions. Nevertheless, it requires an in-depth analysis in which one would optimize customer portfolios incorporating the information available on the correlations. We hope that this research opens up new avenues for a better understanding of the correlations between the timing, spending and dropout decisions of customers.

References

Abe, M. (2009a). ”Counting Your Customers” One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model. *Marketing Science*, *28* (3), 541-553.

Abe, M. (2009b). Customer Lifetime Value and RFM Data: Accounting Your Customers: One by One. *No. CIRJE-F-616. CIRJE, Faculty of Economics, University of Tokyo*.

Ailawadi, K. L., Gedenk, K., Lutzky C. and Neslin S. A. (2007). Decomposition of the Sales Impact of Promotion-Induced Stockpiling. *Journal of Marketing Research. 44* (3), 450-467.

Ailawadi, K. L. and Neslin, S. A. (1998). The Effect of Promotion on Consumption: Buying More and Consuming it Faster. *Journal of Marketing Research, 35* (3), 390-398.

Ainslie, A. and Rossi, P. E., (1998). Similarities in Choice Behavior across Product Categories. *Marketing Science. 17* (2), 91-106.

Baltas, G., Argouslidis, P. C., and Skarmeas, D. (2010). The Role of Customer Factors in Multiple Store Patronage: A Cost-Benefit Approach. *Journal of Retailing. 86* (1), 37-50.

Bell, D. R., and Lattin, J. M. (1998). Behavior and Consumer Preference for Store Price Format: Why “Large Basket” Shoppers Prefer EDLP. *Marketing Science*, *17* (1), 66-88.

Bell, D. R., Chiang, J., and Padmanabhan, V. (1999). The Decomposition of Promotional Response: An Empirical Generalization. *Marketing Science*, *18* (4), 504-526.

Blattberg, R.C., Eppen, G.D., and Lieberman, J. (1981). *A Theoretical and Empirical Evaluation of Price Deals for Consumer Nondurables, in Perspectives on Promotion and Database Marketing*, 101-114. World Scientific Publishing.

Blattberg, R.C., Getz, G. and Thomas, J. (2001). *Customer Equity: Building and Managing Relationships as Valuable Assets*, Boston, Mass.: Harvard Business School Press.

Blattberg R.C., Malthouse E.C., Neslin S.A. (2009). Customer Lifetime Value: Empirical Generalizations and Some Conceptual Questions. *Journal of Interactive Marketing*, *23*, 157-168.

Boatwright, P., Borle, S. and Kadane J. B. (2003). A Model of the Joint Distribution of Purchase Quantity and Timing. *Journal of the American Statistical Association*, *98* (463), 564-572.

Bolton, R. (1998). A 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.

Bolton, R. N., and Lemon, K. N. (1999). A Dynamic Model of Customers Usage of Services: Usage as an Antecedent and Consequence of Satisfaction. *Journal of Marketing Research*, *36* (May), 171-86.

Borle, S., Singh, S., and Jain, D. (2008). Customer Lifetime Value Measurement. *Management Science*, *54* (1), 100-112.

Chandon, P. and Wansink, B. (2002). When Are Stockpiled Products Consumed Faster? A Convenience-Salience Framework of Postpurchase Consumption Incidence and Quantity. *Journal of Marketing Research*, *39* (3), 321-335.

Chintagunta, P. K. (1993). Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households. *Marketing Science*, *12* (2), 184-208.

Chintagunta, P. K., Chu, J., and Cebollada, K. (2012). Quantifying Transaction Costs in Online/Off-line Grocery Channel Choice. *Marketing Science 31* (1), 96-114.

Chintagunta, P. K., and Haldar, S. (1998). Investigating Purchase Timing Behavior in two Related Product Categories. *Journal of Marketing Research*, *35* (1), 43-53.

Danaher, P. J. (2007). Modeling Page Views across Multiple Websites with an Application to Internet Reach and Frequency Prediction. *Marketing Science*, *26* (3), 422-437.

Danaher, P. J., and Hardie, B. G. S. (2005). Bacon with your Eggs? Applications of a New Bivariate Beta-Binomial Distribution. *The American Statistician*, *59* (4), 282-286.

Danaher, P. J., and Smith, M. S. (2011). Modeling Multivariate Distributions using Copulas: Applications in Marketing. *Marketing Science*, *30* (1), 4-21.

Dhar, R. and Wertenbroch, K. (2000). Consumer Choice between Hedonic and Utilitarian Goods. *Journal of Marketing Research, 37* (1), 60-71.

Dubé, J.-P., Hitsch, G. J. and Rossi, P. E. (2010). State Dependence and Alternative Explanations for Consumer Inertia. *The RAND Journal of Economics*, *41* (3), 417-445.

Fader, P. S., and Hardie, B. G. S. (2005). A Note on Deriving the Pareto/NBD Model and Related Expressions. *Working Paper http://www.brucehardie.com/notes/009/*.

Fader, P. S., and Hardie, B. G. S. (2007). Incorporating Time-Invariant Covariates into the Pareto-NBD and BG/NBD Models. *Working Paper http://brucehardie.com/notes/019/*.

Fader, P. S., Hardie, B. G. S., and Lee, K.L. (2005a). RFM and CLV: Using Iso-Value Curves for Customer Base Analysis. *Journal of Marketing Research*, *42* (4), 415-430.

Fader, P. S., Hardie, B. G. S., and Lee, K. L. (2005b). "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., and Shang J. (2010). Customer-Base Analysis in a Discrete-Time Noncontractual Setting. *Marketing Science*, *29* (6), 1086-1108.

Fischer, M., Shin H., Hanssens, D.M. (2013). Brand Performance Volatility from Marketing Spending, *Working Paper*.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions. *Biometrika*, *82* (3), 543-552.

Glady, N., Baesens, B., and Croux, C. (2009). A Modified Pareto/NBD Approach for Predicting Customer Lifetime Value. *Expert Systems with Applications*, *36* (2), 2062-2071.

Gordon, B. R. and Sun B (2013). A Dynamic Model of Rational Addiction: Evaluating Cigarette Taxes, *Working paper*.

Green, W. H. (2003). *Econometric Analysis*. New York: Prentice Hall, 5th edition.

Gupta, S. (1988). Impact of Sales Promotions on When, What, and How Much to Buy. *Journal of Marketing Research*, *25* (4), 342-355.

Gupta, S. (2009). Customer-Based Valuation. *Journal of Interactive Marketing*, *23* (May), 169-178.

Gupta, S., Hanssens, D.M., Hardie, B.G.S., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., and Sriram, S. (2006). Modeling Customer Lifetime Value. *Journal of Service Research*, *9* (2), 139-155.

Gupta, S. and Kim, H.-W. (2010). Value-Driven Internet Shopping: The Mental Accounting Theory Perspective. *Psychology and Marketing, 27* (1), 13-35.

Gupta, S., Lehmann, D. R., and Stuart, J. A. (2004). Valuing Customers. *Journal of Marketing Research*, *41* (1), 7-18.

Hanssens, D. M., and Dekimpe, M. G. (2012). Short-Term and Long-Term Effects of Marketing Strategy. *Handbook of Marketing Strategy*, 457.

Hanssens, D.M., Rust, R.T., Srivastava, R.K. (2009). Marketing Strategy and Wall Street: Nailing Down Marketing's Impact, *Journal of Marketing*, *73* (November), 115-118.

Hartmann, W.R. (2006). Intertemporal Effects of Consumption and Their Implications for Demand Elasticity Estimates. *Quantitative Marketing and Economics*, *4*(4), 325-349.

Hendel, I. and Nevo, A. (2003). The Post-Promotion Dip Puzzle: What do the Data Have to Say? *Quantitative Marketing and Economics*, *1* (4), 409-424.

Howard, J.A., and Sheth, J.N. (1969). *The Theory of Buying Behavior*. John Wiley and Sons Inc., New York.

Jacoby, J. and Kyner, D. B. (1973). Brand Loyalty vs. Repeat Purchasing Behavior. *Journal of Marketing Research*, *10* (February), 1-9.

Jen, L., Chou, C.-H., and Allenby, G. M. (2009). The Importance of Modeling Temporal Dependence of Timing and Quantity in Direct Marketing. *Journal of Marketing Research*, *46* (4), 482-493.

Khan, U., Dhar, R., and Wertenbroch, K. (2005). A Behavioral Decision Theory Perspective on Hedonic and Utilitarian Choice. In S. Ratneshwar and D. G. Mick (Eds.), *Inside Consumption: Frontiers of research on Consumer Motives, Goals, and Desires*. London: Routledge.

Kahn, B. E. and Schmittlein, D. C. (1989). Shopping Trip Behavior: An Empirical Investigation. *Marketing Letters*, *1* (1), 55-69.

Kim, B-D, and Park, K. (1997). Studying Patterns of Consumer's Grocery Shopping Trip. *Journal of Retailing, 73* (4), 501-517.

Kim, B-D, and Rossi P.E. (1994). Purchase Frequency, Sample Selection, and Price Sensitivity: The Heavy-User Bias. *Marketing Letters*, *5* (1), 57-67.

Kumar, V., and Reinartz, W. (2006). *Customer Relationship Management: A Databased Approach*. New York: John Wiley.

Kumar, V., R. Venkatesan, R., Bohling, T., and Beckmann, D. (2008). Practice Prize Report: The Power of CLV: Managing Customer Lifetime Value at IBM. *Marketing Science*, *27* (4), 585-599.

Kumar V., Zhang X., and Luo A. (2014). Modeling Customer Opt-In and Opt-Out in a Permission-Based Marketing Context. *Journal of Marketing Research*, forthcoming.

Krishna, A. (1992). The Normative Impact of Consumer Price Expectations for Multiple Brands on Consumer Purchase Behavior. *Marketing Science*, *11* (3) , 266-28.

Krishna, A. (1994). The Effect of Deal Knowledge on Consumer Purchase Behavior. *Journal of Marketing Research, 31* (1), 76-91.

Krishnamurthi, L. and S.P. Raj (1988). A Model of Brand Choice and Purchase Quantity Price Sensitivities. *Marketing Science*, 7 (Winter), 1-20.

Krishnamurthi, L., Mazumdar, T. and S.P. Raj (1992). Asymmetric Response to Price in Consumer Brand Choice and Purchase Quantity Decisions. *Journal of Consumer Research*, 19 (3), 387-400.

Macé, S., and Neslin, S.A. (2004). The Determinants of Pre- and Postpromotion Dips in Sales of Frequently Purchased Goods. *Journal of Marketing Research*, *41* (3), 339-350.

Manchanda, P., Ansari, A., and Gupta, S. (1999). The Shopping Basket: A Model for Multicategory Purchase Incidence Decisions. *Marketing Science*, *18* (2), 95-114.

McAlister L. (1982). A Dynamic Attribute Satiation Model of Variety-Seeking Behavior. *Journal of Consumer Research, 9* (2), 141-150.

Meade, N., and Islam, T. (2003). Modelling the Dependence between the Times to International Adoption of Two Related Technologies. *Technological Forecasting and Social Change*, *70* (8), 759-778.

Meade, N., and Islam, T. (2010). Using Copulas to Model Repeat Purchase Behavior - An Exploratory Analysis via a Case Study. *European Journal of Operational Research, 200* (3), 908-917.

Meyer, R. J., & Assunçao, J. (1990). The Optimality of Consumer Stockpiling Strategies. *Marketing Science*, 9(1), 18-41.

Nelsen, R. (2006). *An Introduction to Copulas*. New York: Springer.

Park, Y.-H., and Fader, P. S. (2004). Modeling Browsing Behavior at Multiple Websites. *Marketing Science*, *23* (3), 291-314.

Park, S., and Gupta, S. (2012). Handling Endogenous Regressors by Joint Estimation Using Copulas. *Marketing Science*, *31* (4), 567-586.

Pinheiro, J. C., and Bates, D. M. (1996). Unconstrained Parametrizations for Variance Covariance Matrices. *Statistics and Computing*, *6* (3), 289-296.

Rao, R. K. S. and Bharadwaj, N. (2008). Marketing Initiatives, Expected Cash Flows, and Shareholders' Wealth. *Journal of Marketing*, *72* (January), 16-26.

Reinartz, W. J., and 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* (4), 17-35.

Reinartz, W. J., and Kumar, V. (2003). The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing*, *67* (1), 77-99.

Romero, J., van der Lans, R. and Wierenga, B. (2013). A Partially Hidden Markov Model of Customer Dynamics for CLV Measurement. *Journal of Interactive Marketing*, *27* (3), 185-208.

Rust, R. T., Lemon, K. N., and Zeithaml, V. A. (2004). Return on Marketing: Using Customer Equity to Focus Marketing Strategy. *Journal of Marketing*, *68* (1), 109-127.

Sarmanov, O. V. (1966). Generalized Normal Correlation and Two-Dimensional Frechet Classes. *Doklady (Soviet Mathematics)*, *168* , 596-599.
Schmittlein, D. C., Morrison, D. G., and Colombo, R. (1987). Counting your Customers: Who Are They and What Will They Do Next? *Management Science*, *33* (1), 1-24.
Schmittlein, D. C., and Peterson, R. A. (1994). Customer Base Analysis: An Industrial Purchase Process Application. *Marketing Science*, *13* (1), 41-67.
Schweidel, D. A., Fader, P. S., and Bradlow, E. T. (2008). A Bivariate Timing Model of Acquisition and Retention. *Marketing Science*, *27* (5), 829-843.
Schweidel, D. A. and Knox G. (2013). Incorporating Direct Marketing Activity into Latent Attrition Models. *Marketing Science*, 32 (3), 471-87.
Seetharaman, P. B., Ainslie, A. and Chintagunta, P. K. (1999). Investigating Household State Dependence Effects across Categories. *Journal of Marketing Research*, *36* (4), 488-500.
Sklar, A. (1959). Fonctions de Répartition à n Dimensions et leurs Marges. *Publications de l'Institut Statistiques de l'Université de Paris*, *8,* 229-231.
Srinivasan, S., and Hanssens, D. M. (2009). Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions, *Journal of Marketing Research, 46* (3), 293-312.
Sriram, S., Chintagunta, P. K., and Agarwal, M. K. (2010). Investigating Consumer Purchase Behavior in Related Technology Product Categories. *Marketing Science*, *29* (2), 291-314.
Stephen, A. T., and Galak J. (2012). The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Market. *Journal of Marketing Research*, *49* (October), 624-639.
Sun, B. (2005). Promotion Effect on Endogenous Consumption. *Marketing Science, 24* (3), 430-443.
Tarasi, C. O., Bolton, R. N., Gustafsson, A. and Walker, B. A. (2013). Relationship Characteristics and Cash Flow Variability: Implications for Satisfaction, Loyalty, and Customer Portfolio Management, *Journal of Service Research*, *16* (2), 121-137.
Tarasi, C. O., Bolton, R. N., Hutt, M. D. and Walker, B. A. (2011). Balancing Risk and Return in a Customer Portfolio, *Journal of Marketing*. *75* (3), 1-17.
Thomas, J. (2001). A Methodology for Linking Customer Acquisition to Customer Retention. *Journal of Marketing Research*. *38* (2), 262-68.
Urbany, J. E., Dickson, P. R. and Kalapurakal, R. (1996). Price Search in the Retail Grocery Market. *Journal of Marketing, 60* (2), 91-104.
Van Heerde, H. J., Leeflang, P. S. H., and Wittink, D. R. (2004). Decomposing the Sales Promotion Bump with Store Data. *Marketing Science, 23* (3), 317-334.
Venkatesan, R., and Kumar, V. (2004). A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy. *Journal of Marketing*, *68* (4), 106-125.
Venkatesan, R., Kumar, V., and Bohling, T. (2007). Optimal Customer Relationship Management using Bayesian Decision Theory: An Application for Customer Selection. *Journal of Marketing Research*, *44* (November), 579-594.
Vilcassim, N. J. and Jain, D. C. (1991). Modeling Purchase-Timing and Brand-Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity. *Journal of Marketing Research*, *28* (1), 29-41.
Vroegrijk, M., Gijsbrechts, E. and Campo, K. (2013). Close Encounter with the Hard Discounter: A Multiple-Store Shopping Perspective on the Impact of Local Hard-Discounter Entry. *Journal of Marketing Research, 50* (5), 606-626.
Wansink, B., and Deshpande, R. (1994). Out of Sight, Out of Mind: Pantry Stockpiling and Brand-Usage Frequency. *Marketing Letters*, *5* (1), 91-100.

Tables and Figures

Table 1: Literature overview of the correlation between the timing, spending and dropout processes

References (by publication date)	Intra-customer correlation Timing - spending	Inter-customer correlation Timing - spending	Timing - dropout	Spending - dropout	Contractual or non-contractual setting	Industry
Boatwright et al. (2003)	x				Non-contractual	Online grocery
Borle et al. (2008)		x	x	x	Contractual	Direct marketing
Abe (2009a)			x		Non-contractual	Online music retailer, department store, music chain
Abe (2009b)		x	x	x	Non-contractual	Department store, music chain
Glady et al. (2009)	x				Non-contractual	Financial securities
Jen et al. (2009)	x				Non-contractual	B2B (office supply) and B2C (direct marketing in health & beauty)
Fader et al. (2010)			x		Non-contractual	Donation to non-profit
Romero et al. (2013)	x	x			Non-contractual	Online music retailer, hypermarket
Schweidel and Knox (2013)	x	x			Non-contractual	Donation to non-profit
Our study	**x**	**x**	**x**	**x**	**Non- contractual**	**Online music retailer, financial securities, hedonic FMCG, utilitarian FMCG**

Table 2: Summary of possible correlation directions and underlying rationales

Direction/ Rationales	Intra-customer correlation	Inter-customer correlation		
	Timing-spending	**Timing-spending**	**Timing-dropout**	**Spending-dropout**
Positive	***Compensating behavior*** (e.g. a delayed purchase leads to larger quantities)	***High- vs. low-value segments*** (frequent/large buyers vs. infrequent/small buyers)	***Complementary segments*** (frequent/short-life buyers vs. infrequent/long-life buyers)	***Complementary segments*** (large/short-life buyers vs. small/long-life buyers)
Negative	***Purchase accentuation or attenuation*** (e.g. a delayed purchase leads to smaller quantities)	***Complementary segments*** (frequent/small buyers vs. infrequent/large buyers)	***High- vs. low-value segments*** (frequent/long-life buyers vs. infrequent/short-life buyers)	***High- vs. low-value segments*** (large/long-life buyers vs. small/short-life buyers)
Rationales	Price promotion (sensitivity), inventory-based consumption, wear-out in variety seeking, satiation, state dependence, learning effects	Differences in acquisition utility (e.g. preference structure) and transaction utility (e.g. cost of time)	Differences in the strength of the relationship with the firm and in the number of opportunities to dropout	Differences in satisfaction, loyalty and involvement with the firm.

Table 3: Summary statistics of the various data sets

	Mean	Standard deviation	Minimum	Maximum
Online music albums				
Number of repeated transactions	1.04	2.19	0.00	29.00
Time since last transaction (in weeks)	6.85	10.73	0.00	38.43
Average transaction value (in euros)	14.08	25.76	0.00	299.64
Lifetime (in weeks)	32.72	3.33	27.00	38.86
Financial securities				
Number of repeated transactions	4.73	7.24	0.00	48.00
Time since last transaction (in weeks)	9.57	11.14	0.00	35.07
Average transaction value (in euros)	2.38	2.91	0.00	19.73
Age (in years)	48.89	16.18	6.00	99.00
Living area (Dummy, City suburb = 1)	0.43	0.49	0.00	1.00
Current lifetime (in weeks)	34.91	0.78	33.53	36.43
Primary bank (Dummy)	0.90	0.30	0.00	1.00
Utilitarian FMCG				
Number of repeated transactions	7.85	6.65	0.00	19.00
Time since last transaction (in months)	11.69	7.31	0.00	19.00
Average transaction value (in eurocent)	928.64	959.93	0.00	6469
Current lifetime (in months)	16.01	5.21	1.00	20.00
Household size	2.49	1.16	1.00	8.00
Hedonic FMCG				
Number of repeated transactions	2.30	2.47	0.00	12.00
Time since last transaction (in months)	11.40	10.10	0.00	34.00
Average transaction value (in eurocent)	436.75	337.60	0.00	2093.00
Current lifetime (in months)	22.75	8.06	13.00	36.00
Household size	2.34	1.08	1.00	8.00

Table 4: Model fit comparisons

	No correlation	Intra-customer correlation only	Inter-customer correlation only	Intra- & inter-customer correlation
Online music albums		*Frank*	*Gauss*	***Frank & Gauss***
Log-Likelihood	-13,688	-13,660	-13,641	**-13,609**
Number of parameters[11]	7	9	10	**12**
Likelihood ratio t-stat (p-value)	Null model	56 (<0.001)	94 (<0.001)	**158 (<0.001)**
BIC (with n=2,457)[12]	27,431	27,390	27,360	**27,312**
Financial securities		*Gauss*	*Student*	***Gauss & Student***
Log-Likelihood	-67,828	-67,495	-67,149	**-66,815**
Number of parameters	19	25	22	**28**
Likelihood ratio t-stat (p-value)	Null model	666 (<0.001)	1,358 (<0.001)	**2,026 (<0.001)**
BIC (with n=16,434)	135,840	135,233	134,512	**133,902**
Utilitarian FMCG		*Student*	*Gauss*	***Student & Gauss***
Log-Likelihood	-223,974	-219,306	-219,977	**-219,002**
Number of parameters	9	12	12	**15**
Likelihood ratio t-stat (p-value)	Null model	9336 (<0.001)	7994 (<0.001)	**9944(<0.001)**
BIC (with n=23,294)	448,039	438,733	440,075	**438,155**
Hedonic FMCG		*Frank*	*Gauss*	***Frank & Gauss***
Log-Likelihood	-136,484	-135,340	-135,000	**-134,928**
Number of parameters	9	12	12	**15**
Likelihood ratio t-stat (p-value)	Null model	2,288(<0.001)	2,968(<0.001)	**3,112 (<0.001)**
BIC (with n=13,054)	273,053	270,794	270,114	**269,998**

[11] In order to model the intra-customer correlation, at least two parameters are needed (the constant and σ_{intra}). For every customer covariate, one parameter is added (see equation 9). For the inter-customer correlation, three parameters are added (see equation 3).

[12] The number of observations n to calculate the BIC is the total number of transactions made by the customers.

Table 5: Estimated inter-customer correlations

	Timing-spending	Timing-dropout	Dropout-spending
Online music albums			
Sigma (standard error)	0.28*** (0.03)	0.43*** (0.09)	-0.17** (0.08)
Financial securities			
Sigma (standard error)	0.15*** (0.03)	0.67*** (0.05)	0.03 (0.04)
Utilitarian FMCG			
Sigma (standard error)	0.50*** (0.03)	0.13 (0.34)	-0.38** (0.19).
Hedonic FMCG			
Sigma (standard error)	0.76*** (0.03)	0.02 (0.30)	-0.64** (0.30)

***Significant at the 1% probability level, ** Significant at the 5% probability level, * Significant at the 10% probability level

Table 6: Hold-out forecasting comparisons of the models without and with intra- and inter-customer correlation for various customer filters: (i) all customers, (ii) the top 10% customers with the highest estimated intra-customer correlation, (iii) the top 10% customers with the highest actual truncated CLV, and (iv) the customers overlapping in (ii) and (iii).

	(i) All customers	**(ii) Top 10% "intra"**	**(iii) Top 10% " truncated CLV"**	**(iv) Top 10% "intra & truncated CLV"**
Online music albums				
# customers selected	2357	235	235	67
RMSE no correlation	**74.26**	149.01	216.13	247.81
RMSE with correlation	75.09	**145.93**	**213.00**	**244.99**
Financial securities				
# customers selected	3472	347	347	43
RMSE no correlation	**39.15**	32.14	**123.25**	90.27
RMSE with correlation	40.14	**31.91**	126.60	**89.42**
Utilitarian FMCG				
# customers selected	2,968	296	296	65
RMSE no correlation	10,161.52	12,678.18	28,329.71	25,051.32
RMSE with correlation	**9,942.58**	**12,100.40**	**26,148.01**	**23,016.34**
Hedonic FMCG				
# customers selected	5682	568	568	110
RMSE no correlation	2,280.09	2,980.84	6,620.27	7,026.87
RMSE with correlation	**2,139.30**	**2,759.02**	**6,084.71**	**6,450.54**

Table 7: Repartition of repeat customers based on their purchase timing and spending for the utilitarian FMCG dataset

	Infrequent buyers (less than one transaction every 1.90 months)	**Frequent buyers** (more than one transaction every 1.90 months)
Small buyers (less than 7.10 euros per transaction)	<u>**# repeat customers = 951**</u> *Median lifetime* = 18.45 months *Average estimated cash flow while active* = 13.05€/month Positive intra-correlation: 14.4% Non-significant correlation: 81.2% Negative intra-correlation: 4.4%	-
Large buyers (more than 7.10 euros per transaction)	-	<u>**# repeat customers = 951**</u> *Median lifetime* = 29.67 months *Average estimated cash flow while active* = 21.55€/month Positive intra-correlation: 21.9% Non-significant correlation: 70.8% Negative intra-correlation: 7.3%

Figure 1: Transaction values vs. interpurchase times of several purchases made by five imaginary customers, each represented by a different symbol. Customer averages are reported in bold.

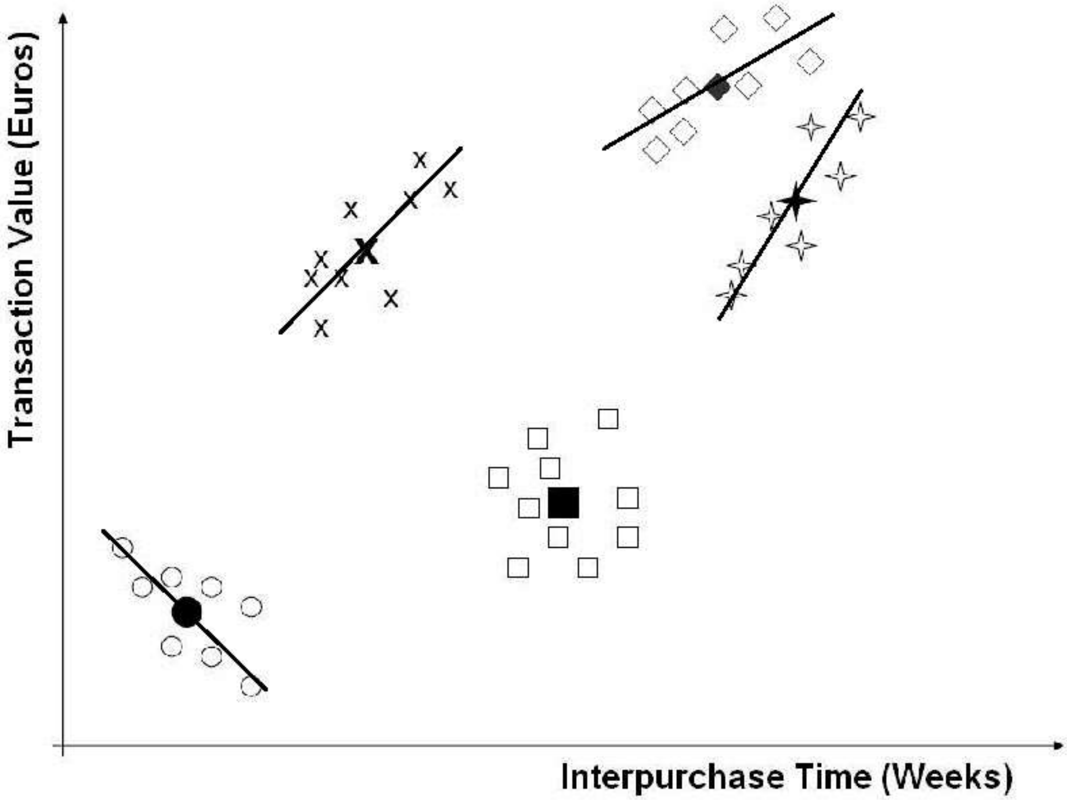


Figure 2: Contours plots corresponding to the joint distribution of two standard-normal random variables with Spearman rank correlation equals to .50, for (i) an independent copula (upper-left plot), (ii) a Gaussian copula (upper-right plot), (iv) a Student copula (middle-left plot) (iv) a Gumbel copula (middle-right plot), and (v) a Frank copula (lower-left plot).

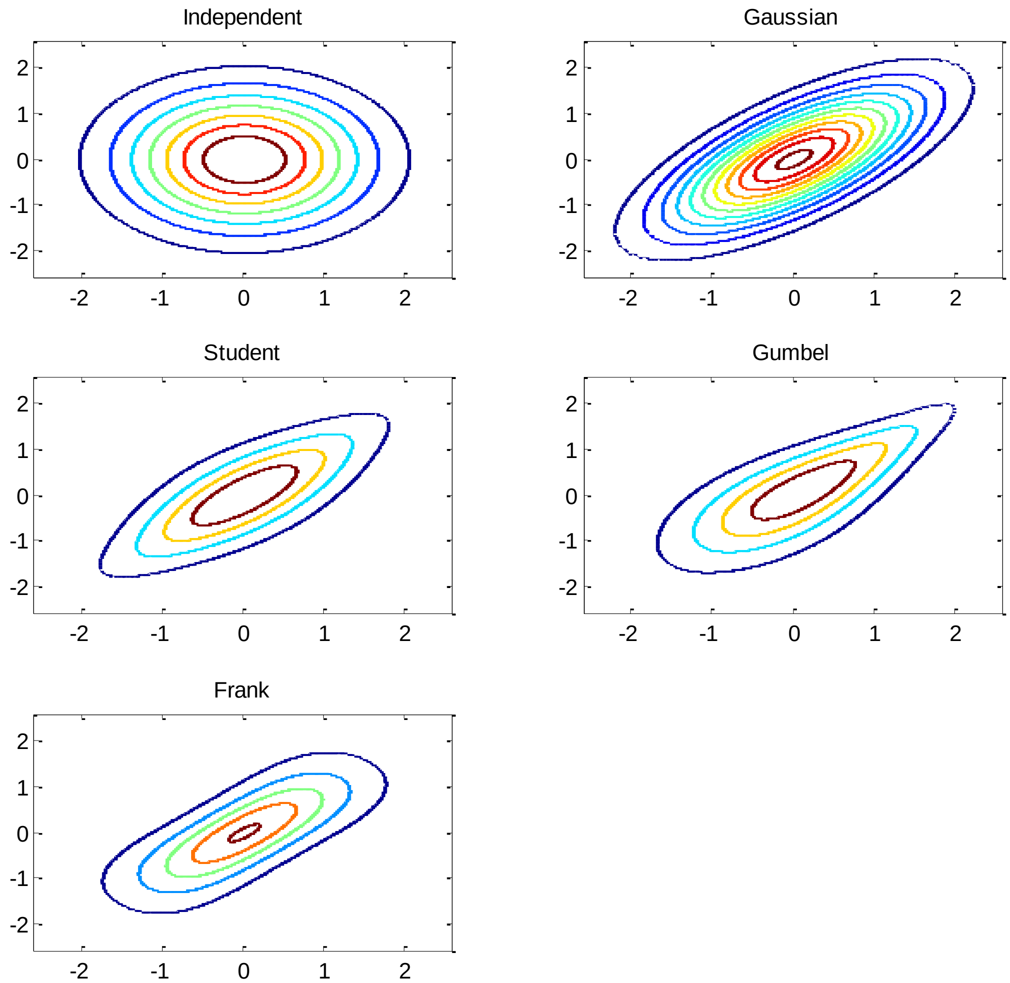
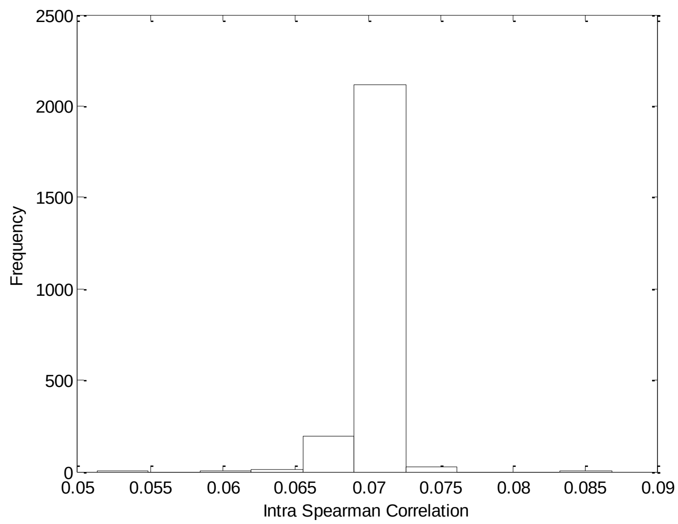
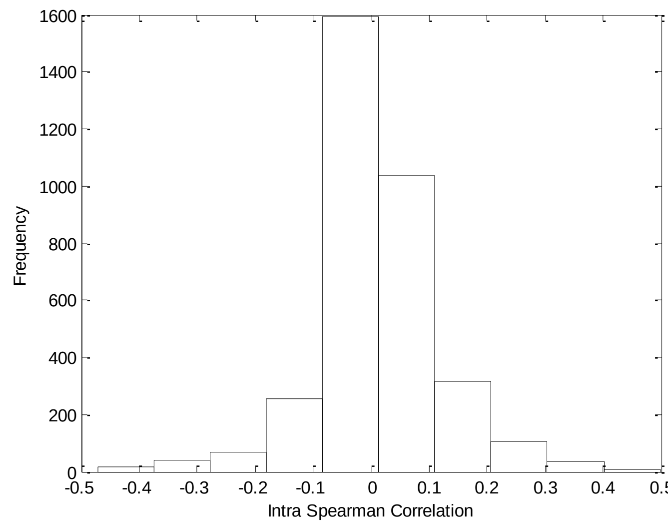


Figure 3: Histogram of the estimated intra-customer correlation for the (a) online music album sales (upper-left panel), (b) securities transactions (upper-right panel), (c) utilitarian FMCG (lower-left panel) and (d) hedonic FMCG (lower-right panel). The x axis reports the spearman correlation.

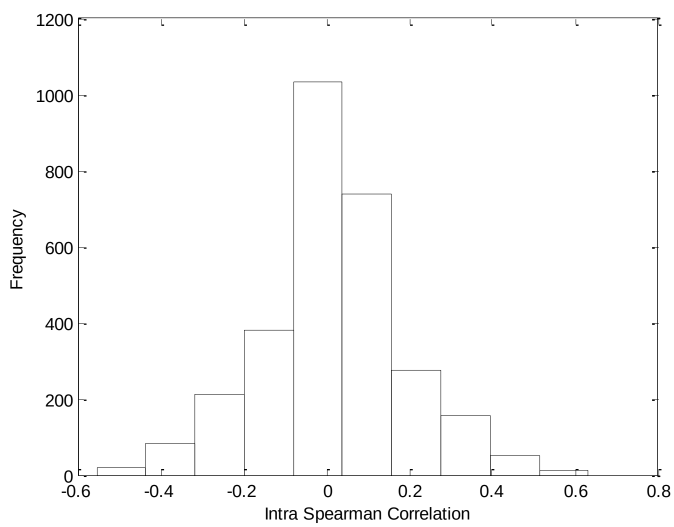
(a) Online music album sales



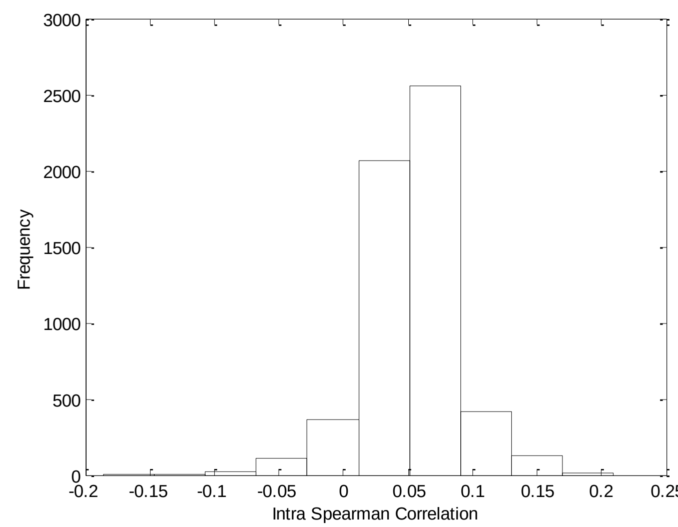
(b) Securities transactions



(c) Utilitarian FMCG



(d) Hedonic FMCG



APPENDIX A: Copula density functions

Here we list the probability density functions of the copula distributions used in this paper.

Gaussian copula A first family of copula that can be found in the literature is the Gaussian or normal copula, with density function

$$c_\theta(F\ \ ,G\ \ ) = \frac{1}{(1-\theta^2)^{1/2}} \exp\left(-\frac{1}{2}\psi'(R(\theta\ \ )^{-1} - I_2)\psi\right), \tag{A.1}$$

where $-1 < \theta\ \ < 1$, $\psi = (\Phi^{-1}(F\ \ ), \Phi^{-1}(G\ \ ))'$, Φ is the univariate standard normal distribution function, $R(\theta)$ is the correlation matrix

$$R(\theta) = \begin{pmatrix} 1 & \theta \\ \theta & 1 \end{pmatrix}, \tag{A.2}$$

and I_2 is the identity matrix of size 2. This copula permits both positive and negative correlation between the variables. Values of θ equal to -1, 0 and 1 correspond to the minimal value of negative correlation, independence, and the maximum of positive correlation. When combined with two normal marginal distributions, the joint distribution is bivariate normal.

Gumbel copula The density of the Gumbel (or logistic) copula is given by

$$c_\theta(F\ \ ,G\ \ ) = \frac{C(F\ \ ,G\ \ )[\log(F\ \ )\log(G\ \ )]^{\theta\ \ -1}}{F\ \ G\ \ [(-\log(F\ \ ))^{\theta\ \ } + (-\log(G\ \ ))^{\theta}]^{2-1/\theta}} \Big[\big((-\log(F\ \ ))^{\theta} + (-\log(G\ \ ))^{\theta\ \ }\big)^{1/\theta\ \ } + \theta\ \ - 1\Big] \tag{A.3}$$

where $1 \leq \theta\ \ < \infty$ is the assocation parameter. The copula distribution in (1) is

$$C(F\ \ ,G\ \ ) = \exp\{-\big((-\log(F\ \ ))^{\theta\ \ } + (-\log(G\ \ ))^{\theta\ \ }\big)^{1/\theta\ \ }\}.$$

The limiting case $\theta\ \ = 1$ gives independence while for θ tending to ∞, one obtains a perfect dependency.

Frank copula The density of the Frank Copula is given by

$$c_\theta(F\ \ ,G\ \ ) = \frac{\theta\ \ (1-e^{-\theta\ \ })e^{-\theta\ \ (F\ \ +G\ \ )}}{[(1-e^{-\theta\ \ }) - (1-e^{-\theta\ \ F\ \ })(1-e^{-\theta\ \ G\ \ })]^2}, \tag{A.4}$$

where $-\infty < \theta\ \ < \infty$. This copula permits both positive and negative correlation between the variables. Values of $-\infty$, 0 and ∞ correspond to the values of smallest negative correlation, independence, and the largest positive correlation respectively.

Trivariate Student copula Let (X_0, Y_0, Z_0) be a trivariate multivariate t-distribution with df degrees of freedom, and correlation matrix R. Denote $F_{X_0}, F_{Y_0}, F_{Z_0}$ the respective marginal distribution. We say that a trivariate random variable (X, Y, Z) has a Student copula with df degrees of freedom and parameter R if the distribution of $(F_X(X), F_Y(Y), F_Z(Z))$ is the same as of $(F_{X_0}(X_0), F_{Y_0}(Y_0), F_{Z_0}(Z_0))$. For $df = \infty$, we find back the trivariate Gaussian copula.

APPENDIX B: The copula-extended CLV model

The model we consider is a two-level model. Conditions A1-A4 below describe the transactions process of an individual customer, i.e. at the intra-customer level. Conditions A5-A9 describe the model at the second level, i.e. at the inter-customer level.

A1: If a customer is alive, the interpurchase times $IPT_{i,j}$ are exponentially distributed with parameter λ_i.

A2: The unobserved time to death τ_i of a customer is exponentially distributed with parameter μ_i.

A3: The transaction values $m_{i,j}$ are gamma(p, ν_i) distributed, with constant shape parameter p, and rate parameter ν_i. In particular $E[m_{i,j}] = p/\nu_i$.

A4: Conditional on the individual parameters λ_i, μ_i and ν_i, we assume that τ_i is independent of the interpurchase times and the transaction values. Furthermore the couples $(IPT_{i,j}, m_{i,j})$ are independent for $1 \le j \le t_{i,x_i}$, but we allow for an correlation between $m_{i,j}$ and $IPT_{i,j}$. This correlation is modeled by a copula c_{intra} depending on a parameter $\theta_{intra,i}$.

A5: Heterogeneity for the transaction rate is modeled as
$\lambda_i = \lambda_{0,i}\exp(\beta_\lambda V_i^\lambda)$.
Here the V_i^λ are the values for the covariates of customer i, and $\lambda_{0,i}$ is a random effect following a gamma(r, α) distribution.

A6: Heterogeneity for the dropout rate is modeled as
$\mu_i = \mu_{0,i}\exp\left(\beta_\mu V_i^\mu\right)$.
Here the V_i^μ are the values for the covariates of customer i, and $\mu_{0,i}$ is a random effect following a gamma(s, β) distribution.

A7: Heterogeneity for the expected transaction value is modeled as
$$E\left[m_{i,j}\right] = \frac{p}{\nu_i} = \frac{p\exp(\beta_\nu V_i^\nu)}{\nu_{0,i}}.$$
Here the V_i^ν are the values for the covariates of customer i, and $\nu_{0,i}$ is a random effect following a gamma(q, γ) distribution. The parameter p is fixed.

A8: The random effects $\lambda_{0,i}, p/\nu_{0,i}$ and $\mu_{0,i}$ follow a trivariate copula c_{inter} with parameter θ_{inter}. Or, conditional on the covariates, the transaction rate, the spending rate and the dropout rate follow a trivariate copula c_{inter} with parameter θ_{inter}.

A9: Heterogeneity for the individual correlation parameter $\theta_{intra,i}$ is modeled as
$f\left(\theta_{intra,i}\right) = \beta_{intra}V_i + \varepsilon_i$, for a suitable link function f and with ε_i following a $N(0, \sigma_{intra}^2)$. The covariates are in the vector V_i, for each customer i.

The (hyper)parameters of this model are in the vector
$$\boldsymbol{\theta} = (r, \alpha, s, \beta, p, q, \gamma, \beta_\lambda, \beta_\mu, \beta_\nu, \theta_{inter}, \beta_{intra}, \sigma_{intra}^2).$$

Conditions A4 and A9 describe the intra-correlation between the spending and the transaction process. Condition A8 deals with the inter-assocation. Conditions A1-A2 are standard for the Pareto/NBD model and conditions A5-A7 introduce the covariates in the model in the same way as in Fader and Hardie (2007). For identification purposes, there is no constant term allowed in $V_i^\lambda, V_i^\tau, V_i^\mu$.

APPENDIX C: Prediction

Below, we outline (1) how we can simulate from the posterior distribution of the individual model parameters, i.e. the transaction rate, spending rate, dropout rate, and intra-customer correlation (2) how the truncated CLV over the next H time units can be predicted from the estimated copula-extended model. We use the notations of Appendix B.

For every customer i, we (1) first generate values from the posterior distribution of the individual parameters, then (2) simulate future transaction data, and compute a value of truncated CLV. By repeating this M times, the distribution of $TCLV_{i,H}$ is simulated. A point estimate is obtained by averaging over the simulated truncated CLV values.

(1) We generate a set of individual parameters $(\lambda_i^*, \mu_i^*, \nu_i^*, \theta_{intra,i}^*)$ from the estimated posterior density $g_i(\lambda, \mu, \nu, \theta_{intra}|data)$. Using the Empirical Bayes principle, we replace unknown (hyper) parameters by their estimates. We sample from the posterior distribution using the independent Metropolis-Hastings algorithm with the estimated prior $g_i(\lambda, \mu, \nu, \theta_{intra})$ as proposal density.

This results in the following iterative scheme:

1. Let $(\tilde{\lambda}_i, \tilde{\mu}_i, \tilde{\nu}_i, \tilde{\theta}_{intra,i})$ be the generated values of the individual parameters in step k of the Markov Chain.
2. Generate $(\lambda_0, \mu_0, \nu_0)$ from the estimated prior distribution specified in assumptions (A5)-(A8). Generate ε_i from a $N(0, \hat{\sigma}_{intra}^2)$, see (A9).
3. Compute
$$\lambda_i = \lambda_0 \exp(\hat{\beta}_\lambda V_i^\lambda)$$
$$\mu_i = \mu_0 \exp(\hat{\beta}_\mu V_i^\mu)$$
$$\nu_i = \nu_0 \exp(-\hat{\beta}_\nu V_i^\nu)$$
$$\theta_{intra,i} = f^{-1}\left(\hat{\beta}_{intra} V_i + \varepsilon_i\right)$$
4. Compute
$$P = \min\left(1, \frac{L_i\left(\lambda_i, \mu_i, \nu_i, \theta_{intra,i}\right)}{L_i\left(\tilde{\lambda}_i, \tilde{\mu}_i, \tilde{\nu}_i, \tilde{\theta}_{intra,i}\right)}\right).$$
The likelihood L_i is given in equation (14). The Markov Chain takes then in step $k+1$ the values $(\lambda_i^*, \mu_i^*, \nu_i^*, \theta_{intra,i}^*)$ with probability P, while with probability $1-P$ the previous values are kept.

(2) For every generated set of individual parameters, we simulate a future transaction stream and the corresponding truncated CLV. First we compute the posterior probability to be alive, $P(\tau_i > T_i|data)$, given the individual parameters, as

$$p_i(\lambda, \mu) = 11 + \frac{\mu}{\lambda + \mu}(\exp((\lambda + \mu)(T_i - t_{x_i})\ \ ) - 1) \tag{A.5}$$

(see Schmittlein et al. 1987). The unobserved time of dropout of customer i, or the time to " death," is denoted by τ_i.

Then we proceed as follows

1. We draw a value from a uniform distribution on [0,1]. If this value is larger than $p_i(\lambda_i^*, \mu_i^*)$, we set $TCLV^* = 0$ and consider the customer as "death." Otherwise, we continue to simulate the transaction process.
2. We draw the time of death τ^* from an exponential distribution with parameter μ_i^*.
3. Set $t^* = 0$ and $TCLV^* = 0$. While $t^* \leq H$ and $t^* \leq \tau^*$
   a) Draw a value (U_1^*, U_2^*) from the copula distribution with parameter $\hat{\theta}_{intra,i}^*$.
   b) Compute IPT^* as the inverse of an exponential cdf with parameter λ_i^* at U_1^*.
   c) Compute m^* as the inverse of the cdf of a $gamma(\hat{p}, \nu_i^*)$ at U_2^*.
   d) Update $t^* \leftarrow t^* + IPT^*$.
   e) Update $TCLV^* \leftarrow \mathrm{T}CLV^* + \frac{m^* \times margin}{(1+d)^{t^*}}$.
   Recall that d stands for the discount rate, and the $margin$ is fixed.

As such, we obtain $M = 4000$ draws from the posterior distribution of the individual parameters and of the $TCLV_{i,H}$. The first 2000 values of the Markov Chain are discarded as burn-in period.

APPENDIX D: Parameter estimates (and standard deviations) for the best-fitting models

	Online music albums		Financial securities		Utilitarian FMCG		Hedonic FMCG	
	No correlation	With correlation	No correlation	With correlation	No correlation	With correlation	No correlation	With correlation
Inter-Correlation								
$\sigma_{tran,spend}$	-	0.28 (0.03)	-	0.15 (0.03)	-	0.50 (0.03)	-	0.76 (0.03)
$\sigma_{tran,drop}$	-	0.43 (0.09)	-	0.67 (0.05)	-	0.13 (0.34)	-	0.02 (0.30)
$\sigma_{drop,spend}$	-	-0.17 (0.08)	-	0.03 (0.04)	-	-0.38 (0.19)	-	-0.64 (0.30)
Intra-Correlation								
Constant	-	0.63 (0.14)	-	-0.07 (0.12)	-	0.09 (0.26)	-	0.48 (0.09)
Household size	-	-	-	-	-	0.18 (0.07)	-	0.40 (0.43)
Living Area	-	-	-	-0.02 (0.08)	-	-	-	-
Lifetime	-	-	-	0.01 (0.07)	-	-	-	-
Primary Bank	-	-	-	0.05 (0.12)	-	-	-	-
Age	-	-	-	-0.38 (0.12)	-	-	-	-
σ_{intra}	-	0.14 (0.26)	-	1.85 (0.07)	-	1.94 (0.15)	-	1.35 (0.17)
Transaction Process								
r	0.55 (0.04)	0.46(0.04)	1.18 (0.04)	1.04 (0.05)	3.60 (0.18)	3.44 (0.66)	1.33 (0.04)	1.16 (0.04)
α	10.58 (0.74)	7.32 (0.59)	2.35 (0.16)	1.84 (0.16)	6.84 (0.33)	6.51 (1.31)	11.22 (0.37)	9.58 (0.37)
Household size	-	-	-	-	-0.32 (0.06)	-0.24 (0.19)	-0.27 (0.06)	-0.27 (0.10)
Living Area	-	-	0.03 (0.04)	0.06 (0.04)	-	-	-	-
Lifetime	-	-	0.50 (0.04)	0.33 (0.04)	-	-	-	-
Primary Bank	-	-	-0.30 (0.07)	-0.20 (0.07)	-	-	-	-
Age	-	-	0.03 (0.05)	0.03 (0.07)	-	-	-	-
Dropout Process								
s	0.61 (0.04)	0.22 (0.04)	2.10 (0.29)	0.97 (0.08)	0.02 (0.00)	0.01 (0.07)	38.91 (3.22)	28.76 (3.62)
β	11.69 (1.19)	4.30 (1.69)	18.04 (3.93)	6.89 (1.20)	0.00 (0.00)	0.00 (0.01)	4988.73 (7.08)	4984.32 (9.89)
Household size	-	-	-	-	0.35 (0.12)	0.73 (0.83)	-0.60 (0.22)	-1.20 (0.57)
Living Area	-	-	0.10 (0.05)	0.16 (0.05)	-	-	-	-
Lifetime	-	-	-0.40 (0.06)	-0.56 (0.06)	-	-	-	-
Primary Bank	-	-	-0.27 (0.10)	-0.24 (0.10)	-	-	-	-
Age	-	-	-0.47 (0.09)	-0.56 (0.09)	-	-	-	-
Spending Process								
p	6.25 (1.19)	6.55 (1.10)	1.20 (0.06)	1.73 (0.04)	0.77 (0.06)	1.48 (0.02)	3.06 (0.18)	1.71 (0.02)
q	3.74 (0.27)	3.31 (0.16)	3.13 (0.13)	2.60 (0.09)	2.85 (0.14)	2.02 (0.06)	4.05 (0.18)	5.05 (0.22)
γ	15.44 (4.14)	11.54 (2.63)	8.15 (0.92)	4.05 (0.20)	2654.83 (322.55)	754.64 (62.75)	510.05 (55.22)	1066.48 (61.93)
Household size	-	-	-	-	-0.46 (0.07)	-0.44 (0.26)	-0.03 (0.06)	-0.05 (0.06)
Living Area	-	-	-0.03 (0.03)	-0.04 (0.04)	-	-	-	-
Lifetime	-	-	0.01 (0.03)	0.04 (0.04)	-	-	-	-
Primary Bank	-	-	-0.38 (0.05)	-0.33 (0.04)	-	-	-	-
Age	-	-	0.82 (0.05)	0.82 (0.04)	-	-	-	-