Intern. J. of Research in Marketing 32 (2015) 78–93

Contents lists available at ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Unveiling the relationship between the transaction timing, spending and dropout behavior of customers[☆]

Nicolas Glady [a,*,1], Aurélie Lemmens [b,1], Christophe Croux [c,1]

[a] Department of Marketing, ESSEC Business School, 1 avenue Bernard Hirsch CS 50 105 95021 Cergy Pontoise Cedex, France
[b] Tilburg School of Economics and Management, Tilburg University, the Netherlands
[c] Faculty of Economics and Business, K.U. Leuven, Belgium

ARTICLE INFO

Article history:
First received on January 17, 2012 and was under review for 6 months
Available online 12 October 2014

Area Editor: Bruce G.S. Hardie
Guest Editor: Hubert Gatignon

Keywords:
Correlation
Copula
Customer lifetime value
Customer portfolio management
Pareto/NBD model
Volatility of cash flows

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.

1. Introduction

Over the last decade, customer lifetime value (hereafter, CLV) has become a powerful customer valuation metric (Blattberg, Malthouse, & Neslin, 2009; Gupta, Lehmann, & Stuart, 2004; Gupta et al., 2006; Kumar & Reinartz, 2006; Kumar, Venkatesan, Bohling, & Beckmann, 2008; Rust, Lemon, & Zeithaml, 2004; Venkatesan & 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 & Reinartz, 2006; Venkatesan, Kumar, & Bohling, 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, Hardie, & Lee, 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 & 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

☆ All authors contributed equally to this paper. This work is part of the Ph.D. thesis of the first author. The second author received financial support from the N.W.O. in the framework of the VIDI grant (project number 016.135.375). The third author received financial support from the FWO research foundation Flanders (research grant G044508N).

* Corresponding author at: Department of Marketing, ESSEC Business School, Avenue Bernard Hirsch, BP 50105, 95021, Cergy-Pontoise, France. Tel.: +33 1 34 43 28 95.
E-mail address: Glady@essec.edu (N. Glady).

[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*) × the price per unit. For simplicity, we will refer to the notion of quantity per transaction or value per transaction interchangeably.

http://dx.doi.org/10.1016/j.ijresmar.2014.09.005
0167-8116/

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 purchase interrelates with the value of this transaction. For instance, Jen, Chou, and Allenby (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, Getz, and Thomas (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, Borle, & Kadane, 2003; Glady, Baesens, & Croux, 2009; Jen et al., 2009; Romero, van der Lans, & Wierenga, 2013), or the inter-customer correlation between the timing, spending and/or dropout decisions (Abe, 2009a,b; Borle, Singh, & Jain, 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 & 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 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 & Hardie, 2005; Danaher & 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.

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

2.1.1. 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.

2.1.2. 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.

2.1.3. 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.

2.1.4. 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

[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).

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