Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 164 (2005) 252–268

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Interfaces with Other Disciplines

Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting

Wouter Buckinx, Dirk Van den Poel *

Department of Marketing, Ghent University, Hoveniersberg 24, 9000 Ghent, Belgium

Received 20 December 2002; accepted 4 December 2003
Available online 27 February 2004

Abstract

Customer relationship management (CRM) enjoys increasing attention as a countermeasure to switching behaviour of customers. Because foregone profits of (partially) defected customers can be significant, an increase of the retention rate can be very profitable. In this paper we focus on the treatment of a company's most behaviourally loyal customers in a non-contractual setting. We build a model in order to predict *partial* defection by behaviourally loyal clients using three classification techniques: Logistic regression, automatic relevance determination (ARD) Neural Networks and Random Forests. Focusing on partial attrition of high-frequency shoppers who exhibit a regular visit pattern may overcome the problem of unidentifiability of total defection in non-contractual settings. Classification accuracy (PCC) and area under the receiver operating characteristic curve (AUC) are used to evaluate classifier performance on a test/hold-out sample. Using real-life data from an FMCG retailer, we show that future *partial* defection can be successfully predicted, i.e. exceeding the benchmark hurdle of the null model. There are no significant differences in terms of performance among alternative classification techniques. Similar to direct-marketing applications we find that past behavioural variables, more specifically RFM variables (recency, frequency, and monetary value) are the best predictors of partial customer defection. This set of variables complements demographic variables confirming findings by other authors about its importance in predicting churn behaviour. Moreover, additional variables (listed in decreasing order of importance) such as the length of customer relationship, mode of payment, buying behaviour across categories, usage of promotions and brand purchase behaviour are shown to be moderately useful to incorporate in attrition models.

Keywords: Marketing; Customer relationship management; Forecasting; Churn analysis; Retention analysis; Retailing; Classification; Logistic regression; ARD neural networks; Random forests

1. Introduction

Customers' life cycles are becoming increasingly transitory due to the severe impact of competitors' actions on existing relationships (Reinartz and Kumar, 2000). Nowadays, consumers are offered a

* Corresponding author. Tel.: +32-9-264-89-80; fax: +32-9-264-42-79.

E-mail addresses: wouter.buckinx@ugent.be (W. Buckinx), dirk.vandenpoel@ugent.be (D. Van den Poel).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2003.12.010

tremendous array of choices. Some people restrict their choices, become relationship oriented (Sheth and Parvatiyar, 1995) and have the potential to become long-life customers. Others exhibit switching behaviour in their shopping (Peterson, 1995). Typically, customers split their purchases among several competitive companies (Dwyer, 1997). This may be due to the fact that customers do not experience any switching costs when changing their supplier (Reinartz and Kumar, 2000).

A relationship has the potential to continue only if both parties are satisfied in the normal setting where alternatives are available (Hoekstra and Huizingh, 1999). If customer satisfaction declines for some reason and a competitor is able to offer a similar product or service, the relationship is likely to be broken. Satisfaction and attractiveness of alternatives determine the strength of relationships (Anderson et al., 1994; Morgan and Hunt, 1994; Peelen et al., 1989). So customer retention is driven by customer satisfaction (as well as other drivers) if sufficient valid alternatives exist (Rust and Zahorik, 1993). Lindgreen and Pels (2002) emphasise that this topic should be studied from a customer's as well as a supplier's perspective. Even if companies are well-equipped to offer relational interaction, some customers prefer not to engage in relationships, i.e. they opt for the 'transactional' exchange as opposed to the 'relational' exchange.

A non-contractual setting suffers from the problem that customers have the opportunity to continuously change their purchase behaviour without informing the company about it. More specifically, in a grocery retail environment (the setting of this study) competition is severe and customers have a wide array of alternatives. This is illustrated by AC Nielsen's (2001) report that more than 70% of all customers shop around in several supermarkets during a month.

Profits can increase because of several reasons (Reichheld, 1996). First of all, by implementing retention programmes, customers are confronted with increasing switching costs, giving them fewer incentives to change their current behaviour (Jones et al., 2000). Secondly, the length of customer relationships influences a firm's profitability. The longer a customer stays the more he spends at the company. Buyers tend to purchase additional services (products) and are more likely to convince others about the positive value the company offers (word-of-mouth effect). They tend to be less price sensitive (Zeithaml et al., 1996) and exhibit a lower responsiveness to competitive pull (Stum and Thiry, 1991).

Retained customers produce higher revenues and margin than new customers (Reichheld and Sasser, 1990). The net return on investments for retention strategies is higher than for acquisitions. So it is supported that companies first spend their marketing resources to keep existing customers rather than to attract new ones (Rust and Zahorik, 1993; Mozer et al., 2000). Recently, however, the argument that customers who purchase steadily from a company over time are necessarily cheaper to serve (or less price sensitive) has drawn substantial criticism (Reinartz and Kumar, 2002).

In summary, customer retention is a valuable strategy to ensure long-term profitability and success of the company. This is illustrated in Table 1. Reducing customer defection can have an enormous impact on companies' results (Mozer et al., 2000; Van den Poel and Larivière, 2004). Suppose 25% of the top clients defect. Considering an average contribution of 2000 Euro a year and a discount rate of 5%, an improvement of the retention rate by just 1% point will cause an increase in profits by 102,923 Euro over five years per 1000 clients (see last column of Table 1).

Table 1
Profit implications

Retention rate	Number of customers left				Total contribution over 5 years (in Euro)	Additional contribution over 75% (in Euro)
	Year 1	Year 2	...	Year 5		
75%	1000	750	...	316	5,679,709	0
76%	1000	760	...	333	5,782,632	102,923

On top of the lost sales new customers need to be attracted which requires very costly actions (Athanassopoulos, 2000; Colgate et al., 1996). Advertising efforts as well as promotions and sales costs are significant but necessary expenses to fill up the customer base (Zeithaml et al., 1996) and establish new relationships (Athanassopoulos, 2000). Besides, new clients often are not profitable for some time.

Moreover, defecting (dissatisfied) customers are convinced that the company offers inferior value and might persuade other customers by spreading negative word-of-mouth (Reichheld, 1996; Sonnenberg, 1990; Mizerski, 1982).

In conclusion, retaining customers by avoiding defection is an important issue for marketing/customer relationship management, CRM managers. A first step in addressing this issue is finding out who to target in retention actions. A fortiori, this is an underresearched topic in the fast-moving consumer goods retail sector. One possible answer is those customers who are most likely to partially attrite. Therefore, in this paper, we investigate whether we are able to predict, at the level of the individual customer, who is going to partially defect. More specifically, we want to find out which of the currently behaviourally loyal customers are likely to (partially) churn in the future. Moreover, we want to gain insight into which predictors are important in identifying partial defection.

This paper is organised as follows. Section 2 presents an overview of the existing body of literature about churn analysis. Section 3 specifies the methodology including three classification techniques used in this study. The description of the data set, as well as an overview of the attributes used to predict customer attrition is discussed in Section 4. Section 5 presents the results and Section 6 phrases the conclusions. Sections 7 and 8 end this paper with a discussion and limitations of this research.

2. Defection of behaviourally loyal customers: Literature review

The topic of customer defection has been discussed extensively in recent literature (cf. Table 2).

Table 2
Literature review

	Sector							Defection		Customer base	
	Retail	Finance	Telecom	Computer manufacturer	Insurance	Automotive firm	Other service	Total	Partial	Complete	Partial
Athanassopoulos (2000)		x						x		x	
Bhattacharya (1998)								x		x	
Keaveney and Parthasarathy (2001)							x	x		x	
Lemon et al. (2002)							x	x		x	
Mittal and Kamakura (2001)						x		x		x	
Mozer et al. (2000)			x				x	x		x	
Popkowski et al. (2000)								x		x	
Van den Poel and Larivière (2004)		x						x		x	
Weerahandi and Moitra (1995)			x					x		x	
Zeithaml et al. (1996)	x			x	x			x		x	
This study	x								x		x

Churn analysis typically tries to define predictors of customer defection. In all of the cases, however, switching behaviour is defined as *total* defection. Customers close their accounts (banking) or change their (mobile) phone operator (telecommunications). In these industries it is easy to observe when defection occurs: people totally interrupt their relationship with the company. As these companies are in a contractual setting, they are able to determine the exact point in time when clients interrupt their relationship. In other sectors it is more complex to determine when customers are leaving. However, buyers typically do not defect from the company all off a sudden. They switch some of their purchases to another store, i.e. they exhibit partial defection. There is a real danger that after a while they will switch completely to the competitor. So in the long run partial defection may lead to total defection.

Table 2 reveals that the churn issue has been underresearched in the retail sector. Moreover, all analyses consider total defection. To discover partial defection this study uses company-internal customer data to determine changes in the individual transaction pattern. We may, for example, hypothesise that customers staying true to their existing patterns are likely to stay, whereas deviations in transaction patterns may signal (partial) defection.

Efforts do not need to be made for the entire customer base. Some customers are not worth the effort to develop a long-term relationship (Hoekstra and Huizingh, 1999). Strategies should be in line with the relationship potential of each customer individually (Reichheld, 1996). It is a well-known phenomenon that a small percentage of customers accounts for a large percentage of profits (Niraj et al., 2001). Moreover, a significant part of the customer base is even not profitable. A small example might illustrate these statements. Imagine a company confronted with a defection rate of 25%. In order to set appropriate marketing strategies, they want to discover why customers defect. A churn analysis for their entire customer base shows that people leave because of the absence of fast checkouts (e.g., cash registers only available to customers who bought less than 10 products). Subsequently, the company decides to invest in such a costly service so more cashiers need to be present at the same moment. However, their most profitable clients are not served with this measure because they typically have more products in their baskets. So only the less profitable customers are satisfied, resulting in a decline of the defection rate, but not necessarily in an increase in profit. In this case, management addressed a reason of customer defection for the unprofitable part of the customer base.

Therefore it is suggested to only focus on those customers in the client base whose future contribution looks promising (Ganesh et al., 2000). Table 2 (last column) reveals that no prior research focused only on the most relevant part of the customer base (in terms of profitability). Instead, they considered all clients.

3. Methodology

3.1. Behaviourally loyal clients

As argued in the previous section, we do not focus on the entire customer base. We only select the best customers of the company. The core of a valuable customer base consists of loyal customers (Ganesh et al., 2000). Loyal customers are more profitable in the short run as well as in the long run (O'Brien and Jones, 1995). They ensure a continuous stream of profits. In our case we focus our study on those who shop frequently and at the same time exhibit a regular buying pattern. To define that segment of clients we use two behavioural attributes: the frequency of purchases and the time between their purchases (interpurchase time or IPT). Both variables are commonly used to define good customers (O'Brien and Jones, 1995). More specifically, the customers in our segment of attention satisfy the following conditions:

1. Frequency of purchases is above average.
2. Ratio of the standard deviation of the interpurchase time to the mean interpurchase time is below average.

The first criterion provides an indication of a customer's loyalty (Wu and Chen, 2000) and potential profitability. The second attribute ensures that the time between customer visits is regular. To identify behaviourally loyal customers, we do not take into account any monetary condition. This is to avoid missing those buyers who do not yet belong to the segment of currently profitable customers but do have a high *potential* value (Niraj et al., 2001).

3.2. Partial defectors

One of the deliverables of this research is an individual-level prediction of the probability to partially defect in the future. In other words, at some specific point in time we want to determine which behaviourally loyal clients in our database may partially switch their purchases to another store (as indicated by 'P' in Fig. 1). So, ultimately, for each individual we need to make an unambiguous conclusion about his future behaviour. As a result, the models we build will be all binary classification models where the dependent variable classifies a particular customer either as a partial defector or as a customer continuing his loyal buying pattern.

However, in a non-contractual setting it is not clear when people defect. Therefore, it is very important to clearly define the concept of partial defection. To this end, we again take into consideration both conditions of the previous paragraph that are used to define our segment of interest but this time over a period of observation *after* the period used to determine behavioural loyalty (i.e., after point 'P' in Fig. 1). So, if one of the above mentioned conditions (1) or (2) is not fulfilled, we classify a customer as partially defective (as the dependent variable) because he deviates from his established transaction pattern.

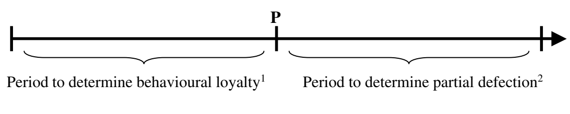

Fig. 1. Period of observation. [1]This period of five months (from April 2000 to August 2000) was also used to derive the independent variables of the model (See Empirical Study). [2]This period of five months (from September 2000 to January 2001) was used to derive the dependent variable.

3.3. Classification techniques

The problem of separating behaviourally loyal customers from behaviourally non-loyal clients may be solved by any classification technique. In this section we discuss the three techniques we use for this task.

3.3.1. Logistic regression

Logistic regression modeling is a well-known technique. It is very appealing because: (1) A closed-form solution for the posterior probabilities is available (as opposed to probit); (2) The basic assumption of logit (the logarithm of the ratio of group-conditional densities is linear in the parameters) is satisfied by many families of distributions (Anderson, 1982); (3) It is easy to use and provides quick and robust results.

In this study we include the technique as a benchmark to compare the more advanced techniques against. We refer to other texts for more technical details (Anderson, 1982).

3.3.2. Automatic relevance determination (ARD) neural network

Artificial neural networks are often credited for achieving higher predictive performance compared to other (statistical) classification techniques (Baesens et al., 2002; Viaene et al., 2002). Within the broad group of neural network architectures we select MacKay's Bayesian ARD neural network framework because it has the appealing property of providing a Bayesian hyperparameter per input variable, representing the importance of the variable (MacKay, 1992). More specifically, we use Nabney's (2001) MATLAB implementation for ARD neural networks. When fixing the number of hidden units, we take into account Penny and Roberts' (1999) recommendation to use a sufficiently large number of hidden units to ensure obtaining a reliable estimate of the predictors' importance.

3.3.3. Random forests

Decision trees have become very popular for solving classification tasks because they can deal with predictors measured at different measurement levels (including nominal variables) and because of their ease of use and interpretability (Duda et al., 2001, Chapter 8). However, they also have their disadvantages such as lack of robustness and suboptimal performance (Dudoit et al., 2002). Recently, many of these disadvantages have been dealt with by creating an ensemble of trees and letting them vote for the most popular class, labelled forests (Breiman, 2001). Several successful paths have been explored how to grow ensembles of trees: (1) Bagging, where to grow each tree a random selection (without replacement) is made from the examples in the training set (Breiman, 1996); (2) Random split selection, where at each node the split is selected at random from among the K best splits (Dietterich, 2000); and (3) Random subspace method, which does a random selection of a subset of predictors to grow each tree. In this paper we select the random forests as proposed in Breiman (2001) which uses the latter strategy. An interesting by-product of these ensembles of trees is their importance measures for each variable. The only two parameters a user of the technique has to determine are the number of trees to be used and the number of variables to be randomly selected from the available set of variables. In both cases we follow Breiman's recommendation to pick a large number (5000 in this case) for the number of trees to be used, as well as the square root of the number of variables for the latter parameter.

3.4. Evaluation criteria

In order to evaluate the performance of classification techniques we use two criteria: percentage correctly classified (PCC) and area under the receiver operating characteristic curve (AUC). Both measures are commonly used as performance criteria (Mozer et al., 2000; Zhang et al., 2002; Chawla et al., 2002). The PCC compares the 'a posteriori' probability of defection with the true status of the customer. The resulting confusion matrix is used to calculate the accuracy of the models. A disadvantage of this measure is that it is not very robust concerning the chosen cut off value in the 'a posteriori' probabilities (Baesens et al., 2002). The AUC measure takes into account all possible cut off levels. For all these points, it considers the sensitivity (the number of true positives versus the total number of defectors) and the specificity (the number of true negatives versus the total number of non-defectors) of the confusion matrix in a two-dimensional graph, resulting in a ROC curve. The area under this curve can be used to evaluate the predictive accuracy of classification models.

4. Empirical study

4.1. General

For our empirical analysis, one of the largest retailers with worldwide operations offering fast moving consumer goods (FMCG) provided the necessary data. Different purchase occasions could be traced by means of a loyalty card. We refer to Ziliani (2000) for an overview of alternative micromarketing (which also comprises CRM) strategies using loyalty-card data. Over 85% of purchases at this particular retailer are registered by their loyalty card. Specifically, we used individual records of 158,884 customers from April 2000 to January 2001, which represented a random sample from the entire customer base containing millions of customers within one geographic area. Even though a five-month period may seem short, we believe it is adequate since we are dealing with an FMCG retailer with an average interpurchase time of 12 days, which results in an average visit rate of 30 times a year.

The first five-month period of the available data, from April to August, is used to define the retailers' behaviourally loyal customers (see Fig. 1). Consequently, we select 32,371 customers, which we consider to be behaviourally loyal clients. This is 20.37% of the total available customer base. These behaviourally loyal clients visit the retailer each week, which means that their average interpurchase time is only seven days (compared to 12 days for the total customer base). Besides, their spending is a lot higher. The average spending of a customer is 1417 Euro a year, whereas the behaviourally loyal customers spend almost twice as much: 2832 Euro. We randomly separated this group of customers in a training set (16 079 observations) and a test set (16 292 observations). The same procedure is used to determine whether they defected during the subsequent period of five months (from September to January). Applying our partial-defection definition results in 8140 partial defections. This is 25.15% of the clients under investigation.

4.2. Predictors

The available data consists of behavioural information at the level of the individual customer and customer demographics. Prior research already supports the incorporation of these two groups of predictors. Table 3 reveals that a major part of the existing attrition studies focuses on demographics as antecedents of defection.

Using the observed past purchase behaviour and additional customer information we compile 61 variables to predict (partial) churn behaviour. These variables have the advantage of being widely available and have shown to be effective and rich predictors (Schmittlein and Peterson, 1994; Buckinx et al., 2004). Table 4 summaries all behavioural independent variables supported by former research. The number of purchases (Frequency) and the amount of spending (Monetary) are the most popular predictors in other research. The time of the day (of purchase or consumption), the length of the customer-supplier relationship (LoR), buying behaviour across categories (Category), mode of payment (MoP), usage of promotions and brand purchase behaviour are variables rarely used in past research. Our study, however, will take them into account.

The following paragraphs provide a motivation for including each of these variables. An overview of all variables used in this study can be found in Table 5.

Table 3
Predictors of defection in prior research

	Behavioural antecedents	Demographics	Perceptions
Athanassopoulos (2000)		x	x
Bhattacharya (1998)	x	x	
Keaveney and Parthasarathy (2001)	x	x	x
Lemon et al. (2002)	x	x	x
Mittal and Kamakura (2001)		x	x
Mozer et al. (2000)	x	x	
Popkowski et al. (2000)	x	x	
Weerahandi and Moitra (1995)		x	
Zeithaml et al. (1996)			x
This study	x	x	

Table 4
Behavioural predictors of defection in prior research

	Recency	Frequency	Monetary	Timing	LoR	Category	MoP	Complaints	Credit
Bhattacharya (1998)	x	x	x		x				
Keaveney and Parthasarathy (2001)		x	x						
Lemon et al. (2002)		x							
Mozer et al. (2000)		x	x	x	x	x	x	x	x
Popkowski et al. (2000)			x						
This study	x	x	x	x	x	x	x	x	x

Table 5
Predictors used in this study

Variable type	Variable name	Description
Interpurchase time	Recency	Number of days since last shop incidence
	MeanRecency	The average number of days between a customers' shop incidences (IPT)
	StdDevRecency	Standard deviation of the IPT
	CVRecency	Coefficient of variation of recency, i.e., ratio of StdDevRecency to MeanRecency
Frequency	Frequency	Number of shop visits (with purchase)
	rFrequency	Number of shop visits relative to the length of relationship (LoR)
	FreqLastMonth	Number of shop visits during last month
	FreqLastWeek	Number of shop visits during last week
Monetary	Monetary	Total monetary amount of spending
	rMonetary	Total spending relative to the length of relationship (LoR)
	rMajorTrip	Percentage of shop visits with above-average spending
Category	rCat (1–12)	Aggregated relative spending in 12 different categories: prepared meals, chemist's, drinks, food, fruit and vegetables, dairy products, meat, non-food, fish, bakery, wine and alcohol, and self-catering
	Cat 1	Aggregated spending in the self-catering category
	NoCat	Number of categories ever purchased from
Brand	NatBrand	Aggregated relative national brand purchase behaviour
	RetBrand	Aggregated relative retailer's brand purchase behaviour
	LowBrand	Aggregated relative low budget brand purchase behaviour
Length of relationship	LoR	Number of days since first purchase
Timing	MeanTimeOfDay	Average moment in time of shopping
	StdDevTimeOfDay	Standard deviation of meantime
	LastTimeOfDay	Time moment of last store visit
Mode of payment	rMoP (1–6)	Aggregated relative amount of money paid in six different ways: 1. cash, 2. check, 3. lunch-allowance check, 4. in-house vouchers, 5. debit card and 6. credit card
	MoP (1–3)	Aggregated amount of money paid in three of the six different ways: 1. cash, 2. lunch-allowance check, and 3. credit card
	rRetBottles	Aggregated relative value of returned bottles
	RetBottles	Aggregated value of returned bottles
Promotions	FreqPromo	Number of shop incidences where coupon was used
	NoVisitsLastCoupon	Number of visits since a coupon was used for the last time
	MeanMonCoupon	Average monetary value of coupons (per shopping trip)
	LoyPoints	Number of loyalty points earned because of special product purchase
Demographics	hhs (1–4)	Household size: number of members in the household
	Language	Language (labels a different language group)
	Title (1–2)	Title of the person
	RegionCode (1–6)	Postal code region classification
	Pets	Presence of pet(s): no (0)/yes (1)
	DemoMissing	Dummy indicating whether or not demographic information is missing

4.2.1. Interpurchase time and related inputs

We include several variables that are related to the time between customers' shop incidences. First, we include 'Recency', which represents the number of days that passed between the last transaction and the end of our observation period. Customers who recently purchased are more likely to be active than customers who shopped a long time ago (Wu and Chen, 2000). Most previous studies find that the lower the value of recency, the higher the probability that a customer stays loyal. In a non-contractual setting this can be the most important variable to indicate an active or inactive relationship (Reinartz and Kumar, 2000). Secondly, the average interpurchase time (IPT), the standard deviation, and the coefficient of variation (ratio of the standard deviation to the average) are incorporated. The average IPT reflects the recency variable over the entire time period. The standard deviation of the IPT and the coefficient of variation measures the irregularity of the time between purchases. We hypothesise the more irregular the less loyal a customer will be.

4.2.2. Frequency of purchases

The customer's frequency of purchases may be predictive for their future behaviour (Schmittlein and Peterson, 1994) because it is positively related to customers' expected future use (Lemon et al., 2002). The probability that a customer is alive may be measured by the number of purchases (Reinartz and Kumar, 2000). Again, we propose several alternative operationalisations of this type of variable. 'Frequency' is the number of shop visits. Moreover, we use the number of days that a person is already a customer at the retailer to include a 'relative' version of the frequency variable. "FreqLastMonth" and "FreqLastWeek" represent the frequency of purchases during the last month and last week of data respectively. Both variables are included because variables computed over more recent time periods may be (more) important to include as predictors.

4.2.3. Monetary indicators

These indicators represent the amount of money someone has spent at a company. The monetary value of each customer's past purchase behaviour tends to be effective in predicting purchase patterns (Schmittlein and Peterson, 1994) and is used in the literature to determine future patterns. Mozer et al. (2000) included monthly charges and usage to predict subscriber dissatisfaction and improve their retention rate. We incorporate three monetary indicators: 'Monetary' is the accumulated amount of money spent from April to September, 'rMonetary' is the same as 'Monetary' but takes into account the length of the relationship of a customer with the retailer, and 'rMajorTrip' indicates the percentage of purchases that could be classified as a big shopping incidence.

4.2.4. Shopping behaviour across product categories

Defection may occur when customers are not pleased (anymore) with a specific product or service (Mozer et al., 2000; Rust and Zahorik, 1993; Mittal and Lassar, 1998). Possible explanations are that prices are too high or quality of the product or service decreases (compared with competitors). If indeed the price or quality of a (category of) product(s) deteriorates and someone intensively purchases this product (category), the probability of defection increases. Consequently, we include inputs representing the spending in each category of the retailer. Literature supports the use of categorical behaviour (Athanassopoulos, 2000). Verstraeten et al. (2002) found preliminary evidence for the existence of a 'natural' order of product purchases. Customers may start their relationship with the retailer by buying specific products. The start of buying specific products or products from certain categories may be the indicator of a changing loyalty towards the company.

The retailer's product-category taxonomy consists of 12 main categories. If numerous customers defect because of the use of a specific category, our model may indicate that the category-spending variable is a predictor of partial defection.

Besides the monetary version we compute the total number of different categories someone purchases from (NoCat). The number of active products/services might be linked to defection (Mozer et al., 2000). The higher this number the more active someone is.

4.2.5. Brand purchase behaviour

The retailer classifies each product into a brand category: national brand, retailer's own store brand, a private label brand, or a (temporal) exclusive brand. For each of these brands a variable is compiled, representing the relative spending of a customer. First, the arguments we used to support the incorporation of the variables summarising their shopping behaviour by category (cf. previous paragraph) can be repeated here. If a significant part of the retailer's top clients defect because of a problem with some brand, the model may indicate that the brand-spending variable is powerful to predict defection. Consequently, management is able to define tailor-made actions. Secondly, concerning the private label/store brand, it is known that qualitative retailer brands can be a tool to differentiate a store and increase store loyalty (Corstjens and Lal, 2000). So we hypothesise that the higher the spending for the store brand/private label brand of the store, the lower the probability that the consumer will leave.

4.2.6. Length of relationship

Length of relationship (LoR) represents the number of days an individual is shopping at the retailer. Bhattacharya (1998) found that the extent to which a customer is able to identify himself with a company is positively related to the period he is willing to continue this relationship. Anderson and Weitz (1989) confirmed this expectation and indicated that the length of the relationship is positively associated to the perceived future stability of the relationship. Verhoef et al.'s (2002) findings confirm the impact of age of relationship on number of services purchased in an insurance context.

4.2.7. Timing of shopping

People do not shop all at the same time during the day or week. This may lead to service quality differences across several moments of the day. For example, employees may be significantly friendlier at noon because in the morning they suffer from morning mood and in the evening they are very busy because the store is too crowded. Under this assumption, people shopping at noon may experience a higher level of service quality than people shopping at other moments. As a result we include a variable representing the average of all points in time when a customer left the shop (check-out time).

4.2.8. Mode of payment

Customers are offered several possible ways to pay their bill. The use of each of these modes of payment might be useful to classify customers into different segments and consequently might be a predictor for future behaviour. The different modes of payment are: cash, checks, lunch-allowance checks, in-house vouchers, electronic payment and credit cards. The in-house vouchers are distributed by the retailer to reward customers for their loyalty based on the information collected by customer loyalty cards. For example, the intensive use of these vouchers might be predictive for upcoming loyalty. The possession and use of a credit card may indicate that customers like to make use of credit. Literature confirms the use of credit information and rate plans for churn analysis (Mozer et al., 2000). An additional variable in this context is the amount of money subtracted from the bill because of returned empty bottles. People returning their empty bottles to a shop show loyalty and consistency towards the retailer.

4.2.9. Promotional behaviour

Prior literature supports that the degree of competition between stores has increased over time. Due to the increased merchandising and promotional activities of retailers consumers are trained to compare deals across competitors (Kim and Staelin, 1999). Moreover, Bawa and Shoemaker (1987) proved that customers being deal-prone are less brand loyal and less store loyal. For them, the lower prices are the explanation of their purchases. These customers typically do not develop a relationship with one specific company. Consequently, we hypothesise that people being more sensitive to promotions will have a higher probability of store switching and thus defection.

4.2.10. Customer demographics

Table 3 indicates the extensive use of customer demographics in other studies of customer defection. Mittal and Kamakura (2001) show that among other things, gender, number of children in a household as well as area of residence are moderating customer characteristics. Vakratsas (1998) confirmed the moderating role of household size: small households are more deal prone than larger-size households (Buckinx et al., 2004). So we expect these clients to be less loyal to the retailer. Mozer et al. (2000) included an indication of the subscriber's location.

Consequently, we incorporate several demographical predictors available in the retailer's data warehouse: 'hhs1'–'hhs4' are dummies in order to indicate that a household exists out of one till four or more members respectively (0/1). Secondly, 'Title1' and 'Title2' indicate the title of the person who subscribed for the loyalty card of the retailer. 'Language' is a dummy representing the mother language of the household. The dummies 'RegionCode1'–'RegionCode6' contain geographical information of the customer and finally 'Pets' makes a distinction between people having one or more pets at home and people without a pet.

For 10% of the customers (3288) these demographics were not available. Consequently, a dummy 'DemoMissing' is added in order to take this into account. At the same time, this variable may be an indication of the level of trust in the company because giving personal information to a firm may be an indication of involvement and confidence.

5. Results

Results presented in Table 6 lead us to conclude that predicting partial defection of behaviourally loyal customers is a viable strategy: First, PCC performance of 0.8040 for random forests on a test sample (i.e. on cases not used during estimation) should be benchmarked to Morrison (1969) proportional chance criterion [1] of 0.6235 ($= 0.2515^2 + (1 - 0.2515)^2$) or the majority prediction rule of 0.7485 ($= 1 - 0.2515$); and second, AUC performance of 0.8310 (again for random forests on the test sample) exceeds the 0.5 benchmark of the null model.

When comparing the different classification techniques they all offer similar performance. Even though random forests consistently come in on top (without the need to tune different parameters, as was the case for ARD neural networks), its performance is not significantly higher than that of the other techniques. Given the recent nature of random forests, we would like to emphasise the attractiveness of this technique for several reasons: 1. Consistent high performance; 2. We confirm Breiman's (2001) observation that the performance results are very robust such that there is not really a need for splitting the sample into an estimation and test sample (similar to logistic regression but unlike neural networks); 3. No need to tune parameters (with the exception of setting the number of trees and the number of variables to be randomly selected from the total set of predictors); 4. Easy computation of variable importance measures; and 5. Reasonable computing times (if logistic regression serves as a reference, random forests are 300 times more 'expensive', which still compares favorably to the 90 000 times more 'expensive' ARD neural networks).

Table 6
Performance results

	PCC		AUC	
	Train	Test	Train	Test
Logistic regression	0.7999	0.8013	0.8278	0.8280
ARD NN	0.8083	0.8040	0.8394	0.8310
Random forests	0.8001	0.8040	0.8249	0.8319

In Table 7 we report the average normalised importance of each predictor for the Random Forests method (Breiman, 2001). When comparing the importance measures of the predictors, a Pearson (Spearman) correlation coefficient of −0.345 (−0.313) [2] between the random forest and the ARD neural network is obtained. The similarity in the ranking of the importances is confirmed by the fact that six of the top-ten variables are the same. We do not report any measures for logistic regression (e.g. standardised estimates) because most measures are prone to multicollinearity, which was clearly present in the dataset, but which is not a problem if the focus is mainly on prediction.

It is clear from the rankings of variable importance that behavioural variables are much more important than demographics. Nevertheless,

[1] The proportional chance criterion is defined by Morrison (1969) as: $C_{\text{pro}} = \alpha^2 + (1 - \alpha)^2$, whereby α represents the actual proportion of the class to predict (in this case: partial churners).

[2] The negative sign of the correlation coefficients was to be expected because a higher importance in the case of random forests is reflected by a higher importance value, whereas for ARD neural networks, the variance is used as a reflection of importance (Breiman, 2001).

Table 7
Importance of variables

No.	Random forests		ARD neural network	
	AvgNormImp	Name of variable	Variance	Name of variable
1	0.99394	Frequency	7.43	Frequency
2	0.86378	MeanRecency	10.65	rFrequency
3	0.82147	rFrequency	19.76	MeanRecency
4	0.74515	LoR	22.18	FreqLastWeek
5	0.67258	FreqLastMonth	34.19	Monetary
6	0.67179	StdDevRecency	44.78	FreqLastMonth
7	0.61325	Monetary	46.31	rMoP2
8	0.56375	rMonetary	54.91	StdDevRecency
9	0.44454	rMajorTrip	59.79	hhs4
10	0.41757	DemoMissing	63.58	Title2
11	0.37740	CVRecency	69.60	LoR
12	0.32867	MeanMonCoupon	77.23	RegionCode6
13	0.31931	Recency	77.65	pets
14	0.30720	rRetBottles	90.41	DemoMissing
15	0.30140	rMoP1	91.75	MoP3
16	0.29828	NatBrand	92.62	rMonetary
17	0.28375	LastTimeOfDay	107.68	rMoP1
18	0.28134	RetBottles	125.55	Title1
19	0.27849	rCat5	127.45	rCat5
20	0.27821	rMoP5	127.57	CVRecency
21	0.27762	rCat1	128.01	rCat2
22	0.27697	rCat2	140.97	RetBottles
23	0.27234	rCat4	146.83	rMoP3
24	0.27167	rCat3	154.23	MeanMonCoupon
25	0.27005	FreqLastWeek	161.17	Recency
26	0.26011	FreqPromo	163.75	RegionCode1
27	0.25709	RetBrand	169.06	hhs2
28	0.25156	LowBrand	174.32	hhs3
29	0.24946	StdDevTimeOfDay	178.45	LastTimeOfDay
30	0.24301	rMoP6	181.00	Cat1
31	0.24226	rCat9	184.23	rMoP6
32	0.23945	rCat10	185.22	rCat4
33	0.23699	rMoP4	192.02	MoP6
34	0.23070	MeanTimeOfDay	194.83	Language
35	0.23057	rCat8	207.85	RegionCode4
36	0.22004	rCat6	211.36	rMoP5
37	0.20848	MoP6	212.20	hhs1
38	0.20727	rMoP3	229.98	RegionCode3
39	0.20334	NoCat	232.07	rRetBottles
40	0.18849	LoyPoints	239.69	FreqPromo
41	0.18286	rCat7	243.18	rCat9
42	0.17623	NoVisitsLastCoup	256.24	NatBrand
43	0.16442	MoP3	265.92	rCat3
44	0.15445	Cat1	270.60	NoCat
45	0.14548	rMoP2	271.67	rMajorTrip
46	0.12864	RegionCode2	292.72	MeanTimeOfDay
47	0.11382	RegionCode4	298.97	rCat1
48	0.11201	RegionCode6	318.03	rCat10
49	0.11173	RegionCode3	338.39	RegionCode5
50	0.10782	Title2	351.97	LoyPoints
51	0.09840	hhs1	395.68	rMoP4
52	0.09252	Language	406.36	rCat8

(*continued on next page*)

Table 7 (*continued*)

No.	Random forests		ARD neural network	
	AvgNormImp	Name of variable	Variance	Name of variable
53	0.09050	RegionCode5	422.38	NoVisitsLastCoup
54	0.08219	RegionCode1	440.58	rCat7
55	0.07765	hhs4	451.59	rCat6
56	0.07390	Title1	521.14	StdDevTimeOfDay
57	0.05410	hhs3	579.89	rCat12
58	0.02106	rCat12	610.55	LowBrand
59	0.02078	pets	897.72	RetBrand
60	0.01461	hhs2	2027.82	RegionCode2
61	0.00864	rCat11	2153.79	rCat11

the latter category cannot be ignored. A model only using behavioural variables (i.e. excluding demographics) results in an AUC of 0.8224 as compared to 0.8319 (see Table 6) in the case of random forests on the test sample. Even though this difference may seem small, it may still translate in a significant impact on the company's profits (cf. Table 1). It is remarkable that the most important demographics variable is actually 'DemoMissing'. It gives empirical support to the conclusion that a behaviourally loyal customer who is not willing to give personal information to the firm may signal future partial defection.

Moreover, within the group of behavioural variables, we find recency, frequency, and monetary (RFM) variables to be the best predictors for separating behaviourally loyal customers from non/less-loyal clients. RFM variables are well-known predictors from the field of direct marketing (Baesens et al., 2002; Van den Poel, 2003). Nevertheless, other 'signals' of loyalty are similarly important, such as the length of relationship, as well as returning empty bottles (RetBottles, rRetBottles). On the other hand, the purchase of retailer brands (RetBrand), as well as the number of categories (NoCat) and the number of loyalty points (LoyPoints) are not important in predicting partial churn.

6. Conclusions

Our empirical results show that classification models can provide an individual's (partial) defection probability given all the individual data collected by the retailer (behavioural as well as customer demographics). Consequently, we are able to track down future (partial) defectors. For managers this classification is very useful in order to establish new marketing strategies towards the companies' clients.

Moreover, we are capable to track down partial defection in contrast with past research that focused on total defection. This contribution is substantial because of several reasons. First of all, since we consider only behaviourally loyal clients the losses in terms of sales may be significant even if customers defect only partially. The average spending of a behaviourally loyal client is 2832 Euro a year. Even if these clients switch only 10% of their expenditure to another store, the effect on turnover is remarkable. So avoiding this switching behaviour is valuable for the retailer (see Table 1: Additional contribution calculation). Secondly, partial defection can escalate and possibly lead to total defection in the long run. Therefore, being able to signal partial defection as early as possible will result in important returns and may even be of greater importance than predicting total defection. Consequently, marketing managers can define which of their customers do have a significant chance to decrease their loyal behaviour towards the company. So they are able to execute specific marketing actions to these clients in order to prevent them from leaving.

The predictive performance of the different classification techniques is very close both in terms of the area under the receiver operating characteristic curve (AUC), as well as for the percentage correctly classified.

We may conclude that, compared to customer demographics, RFM (behavioural) variables are better in separating behaviourally loyal customers from those who have a tendency to (partially) defect. This is somewhat in contrast to the expectations we formulated based on existing research, which strongly emphasises the explanatory/predictive power of the demographic variables.

7. Discussion

This attrition research is carried out in a non-contractual setting. This environment suffers from the fact that customers can continuously switch between competitors without feedback to the original company. As a result, it is very hard to define the exact moment in time when clients leave the company. This paper, however, solves the problem by introducing the aspect of 'partial' defection. Customers are considered to break their relationship when they interrupt their loyal and stable purchasing pattern that they exhibit during a period of five months. Moreover, this paper contributes to the literature by making use of actual customer behaviour instead of intentions of repurchase. Lemon et al. (2002) and Morwitz et al. (1993) confirm the fact that directly observing the (defective) behaviour reveals greater insights.

This study contributes to the literature by not focusing on the entire customer base. Not all clients deserve to be taken into consideration when establishing a retention programme. This can be illustrated by a quote from Blattberg et al. (2000, p. 70): "the goal of customer retention management is not zero defections. Instead a firm should manage its retention rate and choose retention strategies and tactics that best support its mail focus: optimizing customer equity". Accordingly, this paper only targets customers whose future contribution looks promising. The companies' targets need to be economically valuable so the increase in tenure should be achieved at a lower cost than the enhancement in customer value (Carroll, 1993). Consequently, behaviourally loyal clients were selected from a retailer in fast-mover consumer goods. The frequency of purchase as well as the time between purchases are used to distinguish promising shoppers from others. Both variables give an indication of customers' purchasing pattern in terms of occurrence and regularity.

In this paper we focus on identifying partial defectors. However, additional research is required to investigate the actual reasons of the defective behaviour before defining the content of the retention strategy. In other words, the people classified as future defectors can be used to compose focus groups and conduct one-on-one interviews to determine which attributes most determine satisfaction (Trust and Zahorik, 1993).

Once the causes of defection and appropriate strategies are defined, companies still face the complex problem of effective allocation of resources (Trust and Zahorik, 1993). Even knowing what specific steps must be taken, it is hard to determine how much money to spend in order to increase the retention rate and at the same time increase the firm's profitability. Bolton (1998) argues that each method of assessing investments designed to increase retention should take into account the effect of changes on duration lifetimes and lifetime revenues. Mozer et al. (2000) confirm that incentives should be offered to those clients whose probability is above a certain threshold. The threshold should be computed based on the expected savings, the time horizon of evaluation, and the costs of the incentive(s). So, adapted communication actions are needed for different profiles of behaviourally loyal clients according to their spending and their defection probability. Fortunately, our models can produce these defection probabilities. The only element we are missing to compute the expected savings is the impact of the appropriate marketing actions. Therefore, a real-life experiment with different level actions for future potential partial defectors might be a good follow-up study. This would offer information on the impact of several actions for different levels of defection probabilities.

8. Limitations and issues for further research

This study has several limitations. First of all, results are confined to the retail fast-mover consumer goods (FMCG) sector. To some extent generalisations can be made for all other companies active in a non-contractual setting where defection is difficult to detect.

Demographics as well as past purchase behaviour were used as inputs in the models, based on data from a company-internal data warehouse. However, this predictor list can be extended with customer perceptions in order to increase the performance of the models. Regrettably, this type of data is typically unavailable in data warehouses. Recently, Bloemer et al. (2003) show that customer satisfaction data can provide useful insights into identifying customers 'at risk'. Even though this fact limits our ability to gain theoretical insight into customer behaviour processes, it can be anticipated that obtaining these data by sending out questionnaires would be a very laborious and expensive exercise (the more so for a database containing millions of customers). Moreover, we anticipate that including these variables would not necessarily improve our predictive capability and would introduce other problems such as non-response bias. Therefore, we leave this as an issue for further research.

We used five months of available data to determine the focus group of the study and five months to evaluate partial defectors. It is unclear to what extent this time window restriction affects our conclusions. Whenever more data is available, more space is left to change the time window. Moreover, we would be able in that case to evaluate the defective behaviour over a longer time period. This will give the opportunity to check what happens after a while to people classified as partial defectors. That way, the expected lifetime value of a customer can be verified more precisely and appropriate actions can be better established. Finally, when more data are available we would be able to investigate the optimal timing of conducting the study. In other words: how frequent should the retention model be updated in order to optimise the retention rate of the retailers.

More fundamentally, identifying customers as potential (partial) defectors is just a starting point for the managerial process of retaining these customers. Alternative tactics or strategies can be formulated and should be tested in the field to find out where and how the marginal marketing euro is best spent (Baesens et al., 2004).

Acknowledgements

We gratefully acknowledge the financial support of the Fund for Scientific Research Flanders (Belgium): FWO grant no. 3G005501.

References

Anderson, J.A., 1982. Logistic discrimination. In: Krishnaiah, P.R., Kanal, L.N. (Eds.), Handbook of Statistics, vol. 2, pp. 169–191.

Anderson, E., Weitz, B., 1989. Determinants of continuity in convential industrial channel dyads. Marketing Science 8 (4), 310–323.

Anderson, E.W., Fornell, C., Lehmann, D.R., 1994. Customer satisfaction, market share, and profitability: Findings from Sweden. Journal of Marketing 58 (3), 53–66.

Athanassopoulos, A.D., 2000. Customer satisfaction cues to support market segmentation and explain switching behaviour. Journal of Business Research 47 (3), 191–207.

Baesens, B., Verstraeten, G., Van den Poel, D., Egmont-Petersen, M., Van Kenhove, P., Vanthienen, J., 2004. Bayesian network classifiers for identifying the slope of the customer lifecycle of long-life customers. European Journal of Operational Research 156 (2), 242–257.

Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J., Dedene, G., 2002. Bayesian neural network learning for repeat purchase modelling in direct marketing. European Journal of Operational Research 138 (1), 191–211.

Bawa, K., Shoemaker, R.W., 1987. The coupon-prone customer: Some findings based on purchase behaviour across product classes. Journal of Marketing 51 (4), 99–110.

Bhattacharya, C.B., 1998. When customers are members: Customer retention in paid membership contexts. Journal of the Academy of Marketing Science 26 (1), 31–45.

Blattberg, R.C., Getz, G., Thomas, J.S., 2000. Customer Equity: Building and Managing Relationships as Valuable Assets. Harvard Business School Press. 228 pp.

Bloemer, J.M.M., Brijs, T., Vanhoof, K., Swinnen, G., 2003. Comparing complete and partial classification for identifying customers at risk. International Journal of Research in Marketing 20 (2), 117–216.

Bolton, R.N., 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.

Breiman, L., 1996. Bagging predictors. Machine Learning 24 (2), 123–140.

Breiman, L., 2001. Random forests. Machine Learning 45 (1), 5–32.

Buckinx, W., Moons, E., Van den Poel, D., Wets, G., 2004. Customer-adapted coupon targeting using feature selection. Expert Systems with Applications, in press.

Carroll, P., 1993. Revisiting customer retention. Journal of Retail Banking 15 (1), 7–13.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. Smote: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research 16, 321–357.

Colgate, M., Stewart, K., Kinsella, R., 1996. Customer defection: A study of the student market in Ireland. International Journal of Bank Marketing 14 (3), 23–29.

Corstjens, M., Lal, R., 2000. Building store loyalty through store brands. Journal of Marketing Research 37 (3), 281–291.

Dietterich, T., 2000. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting and randomization. Machine Learning 40 (2), 139–157.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. Wiley, NY.

Dudoit, S., Fridlyand, J., Speed, T.P., 2002. Comparison of discrimination methods for the classification of tumors using gene expression data. Journal of the American Statistical Association 97 (457), 77–87.

Dwyer, R.F., 1997. Customer lifetime valuation to support marketing decision making. Journal of Direct Marketing 11 (4), 6–13.

Ganesh, J., Arnold, M.J., Reynolds, K., 2000. Understanding the customer base of service providers: An examination of the differences between switchers and stayers. Journal of Marketing 64 (3), 65–87.

Hoekstra, J.C., Huizingh, E.K.R.E., 1999. The lifetime value concept in customer-based marketing. Journal of Market Focused Management 3, 257–274.

Jones, M.A., Mothersbaugh, D.L., Beatty, S.E., 2000. Switching barriers and repurchase intentions in services. Journal of Retailing 76 (2), 259–274.

Keaveney, S.M., Parthasarathy, M., 2001. Customer switching behaviour in online services: An exploratory study of the role of selected attitudinal, behavioural, and demographic factors. Journal of the Academy of Marketing Science 29 (4), 374–390.

Kim, S.Y., Staelin, R., 1999. Manufacturer allowances and retailer pass-through rates in a competitive environment. Marketing Science 18 (1), 59–76.

Lemon, K.N., White, T.B., Winer, R.S., 2002. Dynamic customer relationship management: Incorporating future considerations into the service retention decision. Journal of Marketing 66 (1), 1–14.

Lindgreen, A., Pels, J., 2002. Buyer-seller exchange situations: Four empirical cases. Journal of Relationship Marketing 1 (3), 69–93.

MacKay, D.J.C., 1992. The evidence framework applied to classification networks. Neural Computation 4 (5), 720–736.

Mittal, B., Lassar, W.M., 1998. Why do customers switch? The dynamics of satisfaction versus loyalty. Journal of Services Marketing 12 (3), 177–194.

Mittal, V., Kamakura, W.A., 2001. Satisfaction, repurchase intent, and repurchase behaviour: Investigating the moderating effect of customer characteristics. Journal of Marketing Research 38 (1), 131–142.

Mizerski, R.W., 1982. An attribution explanation of the disproportionate influence of unfavorable information. Journal of Consumer Research 9 (3), 301–310.

Morgan, R.M., Hunt, S.D., 1994. The commitment-trust theory of relationship marketing. Journal of Marketing 58 (3), 20–38.

Morrison, D.G., 1969. On the interpretation of discriminant analysis. Journal of Marketing Research 6, 156–163.

Morwitz, V.G., Johnson, E., Schmittlein, D., 1993. Does measuring intent change behaviour. Journal of Consumer Research 20 (1), 46–61.

Mozer, M.C., Wolniewicz, R., Grimes, D.B., Johnson, E., Kaushansky, H., 2000. Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. IEEE Transactions on Neural Networks 11 (3), 690–696.

Nabney, I.T., 2001. NETLAB Algorithms for Pattern Recognition. Springer Verlag, London.

Nielsen, A.C., 2001. Major study to track store switching, Retail World, August.

Niraj, R., Gupta, M., Narasimhan, C., 2001. Customer profitability in a supply chain. Journal of Marketing 65 (3), 1–16.

O'Brien, L., Jones, C., 1995. Do rewards really create loyalty? Harvard Business Review 73 (3), 75–83.

Peelen, E., Ekelmans, C.F.W., Vijn, P., 1989. Direct marketing for establishing the relationships between buyers and sellers. Journal of Direct Marketing 3 (1), 7–14.

Penny, W.D., Roberts, S.J., 1999. Bayesian neural networks for classification: How useful is the evidence framework. Neural Networks 12 (6), 877–892.

Peterson, R.A., 1995. Relationship marketing and the consumer. Journal of the Academy of Marketing Science 23 (4), 278–281.

Popkowski, L.P.T.L., Sinha, A., Timmermans, H.J.P., 2000. Consumer store choice dynamics: An analysis of the competitive market structure for grocery stores. Journal of Retailing 76 (3), 323–345.

Reichheld, F.F., 1996. Learning from customer defections. Harvard Business Review 74 (2), 56–69.

Reichheld, F.F., Sasser, W.E., 1990. Zero defections: Quality comes to services. Harvard Business Review 68 (5), 105–111.

Reinartz, W.J., Kumar, V., 2000. On the profitability of long-life customers in a noncontractual Setting: An empirical investigation and implications for marketing. Journal of Marketing 64 (4), 17–35.

Reinartz, W.J., Kumar, V., 2002. The mismanagement of customer loyalty. Harvard Business Review 80 (7), 86–94.

Rust, R.T., Zahorik, A.J., 1993. Customer satisfaction, customer retention, and market share. Journal of Retailing 69 (2), 193–215.

Schmittlein, D.C., Peterson, R.A., 1994. Customer base analysis: An industrial purchase process application. Marketing Science 13 (1), 41–67.

Sheth, J.N., Parvatiyar, A., 1995. Relationship in consumer markets: Antecedents and consequences. Journal of the Academy of Marketing Science 23 (4), 255–271.

Sonnenberg, F., 1990. Marketing to Win. Harvard Business School Press.

Stum, D., Thiry, A., 1991. Building customer loyalty. Training and Development Journal 45 (4), 34–36.

Vakratsas, D., 1998. Household cost effects on purchase timing decisions: Do demographics matter? Journal of Consumer Marketing 15 (1), 6–22.

Van den Poel, D., 2003. Predicting mail-order repeat buying: Which variables matter? Tijdschrift voor Economie en Management 48 (3), 371–403.

Van den Poel, D., Larivière, B., 2004. Customer attrition analysis for financial services using proportional hazard models. European Journal of Operational Research 157 (1), 196–217.

Verhoef, P.C., Franses, P.H., Hoekstra, J.C., 2002. The effect of relational constructs on customer referrals and number of services purchased from a multiservice provider: Does age of relationship matter? Journal of the Academy of Marketing Science 30 (3), 202–216.

Verstraeten, G., Van den Poel, D., Prinzie, A., Van Kenhove, P., 2002. Detecting sequential patterns for cross-selling fast moving consumer goods. In: Zanasi, A., Brebbia, C.A., Ebecken, N.F.F.E., Melli, P. (Eds.), Data Mining III. WIT Press, Bologna, pp. 521–530.

Viaene, S., Baesens, B., Van Gestel, T., Suykens, J.A.K., Van den Poel, D., Vanthienen, J., De Moor, B., Dedene, G., 2002. Knowledge discovery in a direct marketing case using least squares support vector machines. International Journal of Intelligent Systems 16 (9), 1023–1036.

Weerahandi, S., Moitra, S., 1995. Using survey data to predict adoption and switching for services. Journal of Marketing Research 32 (1), 85–96.

Wu, C., Chen, H.L., 2000. Counting your customers: Compounding customer's in-store decisions, interpurchase time and repurchasing behaviour. European Journal of Operational Research 127 (1), 109–119.

Zhang, Q., Varshney, P.K., Wesel, R.D., 2002. Optimal bi-level quantization of i.i sensor observations for binary hypothesis testing. IEEE Transactions on Information Theory 48 (7), 2105–2111.

Zeithaml, V.A., Berry, L.L., Parasuraman, A., 1996. The behavioural consequences of service quality. Journal of Marketing 60 (2), 31–46.

Ziliani, C., 2000. Retail micro-marketing: Strategic advance or gimmick? The International Review of Retail, Distribution and Consumer Research 10 (4), 355–368.