EJ **Expert Journal of Marketing (2013) 1**, 38-49

Marketing.ExpertJournals.com

Models Used for Measuring Customer Engagement

Mihai ȚICHINDELEAN*

Lucian Blaga University of Sibiu

The purpose of the paper is to define and measure the customer engagement as a forming element of the relationship marketing theory. In the first part of the paper, the authors review the marketing literature regarding the concept of customer engagement and summarize the main models for measuring it. One probability model (Pareto/NBD model) and one parametric model (RFM model) specific for the customer acquisition phase are theoretically detailed. The second part of the paper is an application of the RFM model; the authors demonstrate that there is no statistical significant variation within the clusters formed on two different data sets (training and test set) if the cluster centroids of the training set are used as initial cluster centroids for the second test set.

Keywords: *probability model, parametric model, relationship marketing, Pareto/NBD model, RFM model*

JEL Classsification: *M31, C12, C14, C38*

1. Introduction

The classic marketing perspective – as practice and science – is transaction oriented. Due to several factors, such as the globalization process and highly informed customers, a criticism to this classical view has emerged in the form of relationship marketing. In a narrow perspective, the objective of relationship marketing is closely related to customer relations (Bruhn 2009, Berry, 1983), from attracting and maintaining, to enhancing these relationships (Berry, 1983). The wider view of relationship marketing spans over the relations a company has with the entities (stakeholders) of the network in which it is active part of. All definitions of relationship marketing (Grönroos, 1990, Bruhn, 2009) have four common characteristics: orientation towards the company's stakeholders, orientation towards the management process, time and a focus on needs.

Presumably, the most important stakeholder group is represented by the company's clients. The customers' needs vary by intensity, form and time, therefore a business's relation with a particular customer or group of customers is dynamic and requires special attention. In this way, the consumer is not viewed as a passive recipient of the company's value creation efforts anymore (Bijmolt, 2010), but as an endogenous element for the company who can co-create value and collaborate to design the company's innovative process (Van Doorn et al., 2010). Thus, the customers' efforts are placed on the same business process level with the company's in the value creation.

A framework of this value co-creation is the concept of value-chain developed by Bruhn (2009) as a theoretical part of understanding relationship marketing. Basically, this framework consists of four elements:

* Correspondence:
Mihai Tichindelean, Lucian Blaga University of Sibiu, Faculty of Economic Sciences, mihaitichi@gmail.com

Article History:
Received 15 October 2013 | Accepted 05 November 2013 | Available Online 18 November 2013

Cite Reference:
Tichindelean, M., 2013. Models Used for Measuring Customer Engagement. *Expert Journal of Marketing*, 1(1), pp.38-49

i. The company's input (the company's marketing activities)
ii. Psychological effects specific to the customers
iii. Behavioral effects specific to the customers
iv. The company's output (the company's financial results)

The interaction of these elements make the value chain model similar to a black-box model, where the inputs take the form of the company's input and the outputs, namely the behavioral effects of the customers, determine the company's output. The black-box is represented by the psychological dimensions of the customer, which are not visible, and can only be deduced by the company. Many studies were undertaken with the purpose of getting insights regarding these dimensions and how they interact: quality of the company's performance (Zeithaml, Parasuraman & Berry, 1992), perceived value (Sanchez-Garcia et al., 2007), customer satisfaction (Oliver, 1996), customer trust (Morgan &Hunt, 1994), customer commitment (Lacey, 2007), and the quality of the relationship (Iacobucci, 2001).

This article focuses on the psychological dimension of *customer commitment* or *customer engagement*. It is defined as the customer's ultimate outcome with causal precedence of satisfaction and trust (Morgan/Hunt 1994). An alternative expression for customer commitment is customer engagement which is defined by Van Doorn et al. (2010) as the customer's behavioral manifestation towards a brand or a firm which goes beyond purchase behavior. This behavioral manifestation can be associated with the customers' behavioral effects within the value chain. Based on these two dissimilar points of view, a sensible difference in understanding a customer's commitment can be seized. If commitment is understood as a psychological dimension, then its intensity (formation) is directly linked with other psychological dimensions (such as satisfaction, trust, perceived value) and *customer-related exogenous factors* (company specific – brand reputation). An idea of its measurement can be based on the theory of dynamic systems, where the earlier mentioned elements can be part of such a system. The main limitations regarding these models (persistence models, Gupta 2006) can be grouped into two categories. *First*, longitudinal data is necessary in order to apply these kinds of models. Although, there are several tools (especially online tools) which measure customers' dimensions such as satisfaction, trust, etc., on the long-term, these psychological dimensions are not ranked as accurate as transactional data is. *Second,* it is quite difficult to measure and to analyze, in a correct manner, psychological dimensions within transversal marketing researches (one time); conducting longitudinal studies based on cohorts of customers make these tasks more difficult.

The other point of view understands customer commitment through its behavioral manifestations. Some of these are visible to the company (purchases), other are not (word of mouth, customer cocreation and complaining behavior). Both of them have an influence on the company's outcome or performance, which can be a direct one (purchase (visible) behavior) or indirect, such as Word-Of-Mouth, loyalty (Bruhn, 2010), participation in the company's activities, customer voice or service improvements (Bijmolt, 2010). Thus, the company's performance is linked to customer engagement; the latter one is understood and measured through a company's customer driven actions. Most of the models used to measure customer engagement are based on transactional data of the customers. Transactional data such as: purchase frequency, purchase volume, purchase value, recency of last purchase can easily be obtained and analyzed by the company. Using only such *visible* behavioral manifestation of the customers' engagement may exclude other important factors (WOM, customer cocreation, and complaining behavior – Bijmolt 2010) which can result in flawed-driven insights regarding the influence of the customers' commitment on the company's performance.

The purpose of this paper can be structured into the following two key subsequent aims:

i. To review the current model developments used to measure the customers' commitment
ii. To apply one model (a scoring models – RFM model) on primary data.

Models used for measuring customer engagement can be differentiated according to the stages of the customers' lifecycle: customer acquisition stage, customer retention stage and customer win-back stage (Bruhn, 2010). The paper of Bijmolt, et. all (2010), reviews these models and groups them according to the stages of the customers' lifecycle.

The first stage of the customers' lifecycle – customers' acquisition stage – is important because of the fact that a customer relationship can be initiated, or not; and in the favorable scenario it can be profitable for the company starting with that particular moment. Phases of this initial stage are divided in an initiation phase and a socialization phase (Stauss, 2008). Within the initiation phase, the customer is searching for information regarding the desired product or service. This phase ends and the second one begins after a transaction (purchase) has occurred, thus a relation between the two parts is initiated (Bruhn, 2010). Bijmolt et all divides this first stage from the company's perspective into phases specific for customer selection and customer acquisition management (Table 1). The goal of customer selection is to identify the *right* customers for further allocation of the company's resources. This *right association* does usually comprehend criteria such as: response likelihood, purchase volumes/values, purchase probabilities, purchase frequency and other transactional data. The models used within the selection phase of customers are further discussed in this paper.

Table 1. *Models for measuring customer engagement within the customer acquisition phase*

Type of data	- RFM data - Customer characteristics (e.g. demographics) - Company-interaction variables (e.g. marketing actions) - Clickstream data	
	Customer selection	**Managing customer acquisitions**
Type of methods	*Parametric (scoring) models*	*Parametric models*
	- RFM scoring (Gupta et.al 2006) - CHAID (David Shepard Associates, 1999) - Linear regression model (Bauer, 1988) - Sequential probit model (Sismeiro and Bucklin, 2004) - Latent class probit model (Vroomen, Donkers, Verhoef and Franses, 2005)	- Logit/Probit model (Hansotia and Wang, 1997; Lewis, 2005; Reinartz, Thomas and Kumar, 2005; Verhoef and Donkers, 2005) - Tobit model (Hansotia and Wang, 1997; Lewis, 2006) - Hazard model (Thomas, Blattberg and Fox, 2004) - Generalized gamma model (Venkantesan and Kumar, 2004) - Hierarchical Bayesian model (Ansari and Mela, 2003) - Poisson count model (Andreson and Simester, 2004)
	Semi-/Nonparametric (scoring) models	*Data/web usage mining*
	- Semiparametric logit model (Bult and Wansbeek, 1995; Malthouse, 2001) - Neural networks (Baesens et al. 2002; Malthouse and Blattberg, 2005)	- Transaction/usage clustering (Mobasher, Cooley and Srivastava, 2000) - Association rule discovery (Mobasher, Cooley and Srivastava, 2000; Mobasher et al. 2001) - Fuzzy inference engine (Nasraoui and Petenes, 2003) - UBB Mining (Ting, Kimble and Kudenko, 2005)
	Probability models	
	- Pareto/NBD model (Abe, 2009; Fader, Hardie and Lee, 2005;Schmittlein, Morrison and Colombo, 1987) - BG/NBD model (Fader, Hardie and Lee, 2005; Fader, Hardie and Shang, 2010) - Individual-level probability model (Moe and Fader, 2004)	

2. Models for Customer Engagement within the Customer Selection phase

2.1. Probability models

A probability model is a representation of the studied phenomena in which observed characteristics (happened events) are the result of a stochastic process of underlying, unobserved characteristics which vary in intensity across the sample studied. A probability model is built on logical deductions regarding the interdependence of the identified characteristics of the studied phenomena. An identified characteristic is a variable which *varies* in a stochastic or random manner across the studied individuals. This random variable usually follows a known probability distribution and is defined by either its probability function (discrete random variables)/probability density function (continuous random variables) or by its cumulative distribution function (both discrete and continuous variables).

There are three main approaches which model transactional data with the purpose of customer selection:

i. Pareto/NBD model (Abe 2009; Fader, Hardie and Lee, 2005; Schmittlein, Morrison and Colombo, 1987)
ii. BG/NBD model (Fader, Hardie and Lee, 2005; Fader, Hardie and Shang, 2010)
iii. Individual-level probability model (Moe and Fader, 2004)

In their initial paper (Schmittlein, Morrison and Colombo, 1987), the authors used transactional data to predict future *transactional* behavior of the studied individuals or customers. They used two sets of data – the customer's recency (time of last purchase) and the customer's purchase frequency (number of transactions completed in an observed time-period). Two different sets of information were attracted from each observed customer and used as input data for the Pareto/NBG model: the *number of repeated transactions* each customer has completed in an observed time period (noted as *x*) and the *period of his last transaction* (noted as t_x with the starting time of 0 which represents the beginning of the observed time period). A clear illustration of the meaning of the data over the observed time period is presented in the figure bellow:

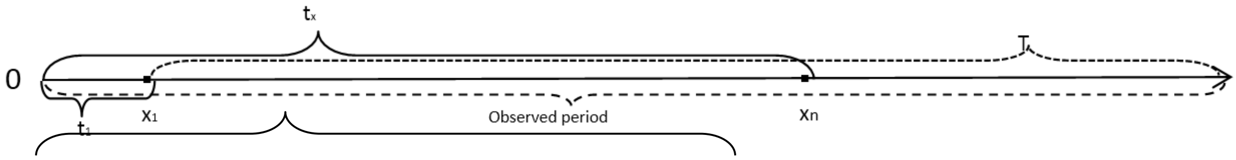

Figure 1. *Graphical representation of the Pareto/NBG model variables*

Every observed customer is symbolized through the notation X = (x, t_x, T). Based on this data and on the timing of the events, the following assumptions were developed for the underlying model (Schmittlein, 1987):

i. A customer's relationship with the company has two phases: he or she is "alive" for an unobserved period of time, and then becomes permanently inactive.
ii. While "alive", the number of transactions made by a customer can be characterized by a Poisson process.
iii. Heterogeneity in the transaction rate across customers follows a gamma distribution.
iv. Each customer's unobserved "lifetime" is distributed exponential.
v. Heterogeneity in dropout rates across customers follows a gamma distribution.
vi. The transaction rates and dropout rates vary independently across customers'.

The *dropout* of any observed customer can be defined as the moment or as the event by which the customer has reached his buying saturation and from which on no further transactions will be completed. The initial SMC (Schmittlein, Morrison, Colombo) Pareto/NBD model assumes the dropout of an observed customer can happen at any moment in time, independent from his actual buying behavior.

The output information provided by the model includes the following:

i. the probability that a customer is still active at a specific moment of time – P(alive| x, t_x , T)
ii. the expected number of transactions that an observed customer will complete in a future period of time (E(Y(t)|x, t_x, T)

A slight variation from the Pareto/NBD model has been developed by the marketing scientists Fader, Hardie and Lee (Counting your Customers the Easy Way: An Alternative to the Pareto/NBD Model - 2005). The assumptions of their probability model are similar to those of the Pareto/NBD model with the exception of the *dropout timing*, which is dependent of the completed transactions, therefore can occur at any point in time *after* a transaction has been completed. The five assumptions can be summarized as follows (Fader, Hardie and Lee, 2005):

i. While active, the number of completed transactions made by a customer follows a Poisson process with an average success value λ. This is equivalent to the assumption that the time between two completed transactions of one customer can be modeled by an exponential distribution with a constant transaction rate λ:

$$f(t_j | t_{j-1}; \lambda) = \lambda e^{-\lambda(t_j - t_{j-1})}, \ t_j > t_{j-1} \geq 0.$$

ii. λ is distributed according to a gamma distribution with parameter α and r

$$f(\lambda | r, \alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda\alpha}}{\Gamma(r)} \quad , \lambda > 0.$$

iii. After each completed transaction, a customer becomes inactive with a probability of p. This point at which a customer becomes inactive is distributed across the transactions according to a (shifted) geometric distribution with a probability function:

$$P(\text{customer becomes inactive immediatly after the } j\text{th transaction}) = p(1-p)^{j-1}$$

$$j = 1, 2, 3, \dots .$$

iv. The probability (p) varies within the customer database according to a beta distribution of parameters a and b with the following probability density function:

$$f(p|a,b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, \ 0 \le p \le 1, \text{ where } B(a,b) \text{ is the beta function.}$$

v. The transaction rate λ and the dropout probability p (probability that the customer becomes inactive after a specific transaction) vary independently one from another across the customer base.

Because of the assumptions *iv.* and *v.* and the underlying distributions (geometric distribution of and beta distribution), the probability model was named BG/NBD model (beta-geometric/negative-binomial-distribution). The likelihood function for every observed customer was constructed and the parameters of the used distribution (of the model specifically) were estimated using optimization methods. As an output, this model offers relevant information regarding: the probability of observing *x* purchases in a period of length t (*P(X(t))*, the expected number of purchases in a time period of a *t* length (*E(X(t)*), and the forecast of the number of purchases an observed customer will complete in a future period of time *t* (*E(Y(t))*).

The two described probability models used for measuring customer engagement within the acquisition phase are solid instruments which provide relevant and easy to understand information. Another strength of these models relies in the small data set needed for application. Two behavioral characteristics of the customers, observed over a period of time, are the only necessary input. Due to longitudinal data, the model can be tested and validated, obtaining valuable forecasts.

2.2. Parametric (scoring) models

Parametric or scoring models can be defined as a set of methods used mainly to group (cluster) the customers according to some grouping variables. Traditional scoring models use behavioral characteristics as grouping variables (recency of last purchase, frequency of purchases within an observed period, monetary value of last purchase) based on the assumption that future buying behavior of the customer is similar to his past behavior. The formed groups or segments of customers prioritize the company's marketing activities according to the scores of the underlying variables. These models describe the customers' commitment through its observed behavioral characteristics (RFM variables), neglecting other possible drivers of it.

The first parametric model which uses RFM variables was proposed by Hughes (1994) with the purpose to differentiate *important* customers from a large database according to the mentioned variables. These variables were defined as: R – Recency of the last purchase (the time interval between the last transaction of the customer and the present time); F – Frequency of purchases (the number of transactions completed in a specific period of time) and M – Monetary value of the purchases (the money spent by the customer for the company's offer in a specific period of time). The importance of these variables is set up through weights. Hughes considers that the three variables are equal in importance, thus they have equal weights (1994). Other authors (Stone, 1995) consider that the weights of these variables should be established according to the researched industry.

Equal or unequal weighted variables are used as input for a clustering procedure. A clustering procedure is an iterative algorithm which groups or *clusters* the observed objects according to their similarities. Thus, the formed clusters differ one from another by the dissimilarities of the underlying objects. Choosing a clustering procedure depends on many factors, among which: the clustering objective, the size of the data and the scales used to measure the input variables are of high importance. There are three commonly used clustering procedures: hierarchical cluster analysis, k-means cluster analysis and two-step cluster.

The scoring model for measuring customer engagement uses the k-means algorithm as clustering procedure. K-means, originally known as Forgy's method (Forgy, 1965), groups the observed objects in *k* *clusters* according to their mean value (an object can be described by one or more variables measured through interval or ratio scales). The steps of this algorithm are:

i. k initial clusters are formed out of the first k objects (the researchers indicates the number of the desired clusters according to his research experience and research objectives). The mean of every cluster is computed (this mean is called a centroid and can be considered a point in an n-dimensional space, where n is the number of variables through which an object is characterized).
ii. The remaining objects are assigned to each of the k clusters according to the smallest distance between the objects' mean and the k clusters' centroids. The distances are computed using Euclidean distances in an n-dimensional space. After all objects have been assigned to a cluster, the centroids of the clusters are computed again.
iii. All objects are classified in one of the k clusters according to the smallest distance between the objects' mean and the new clusters' centroids.
iv. Step iii. is repeated until the clusters' centroids do not exhibit a significant change.

This kind of scoring model based on RFM variables is primary used within the customer acquisition stage, when selecting the *right* customers for future marketing actions is extremely important. Although it has several advantages, the RFM scoring model is considered to have the following limitations (Fader, Hardie, and Lee, 2005; Kumar 2006a): firstly, this model predicts the customers' engagement only in the next period of time, expectations regarding future periods cannot be made; secondly, the used variables (RFM) are observed indicators (observed effects) of the customers' engagement - the formation factors or components are not taken into consideration; thirdly, the model ignores the possibility that the measured customers' engagement is the result of the company's past marketing actions.

2.3. Research methodology

This section of the paper presents a detailed scoring model based on RFM variables and its applicability in segmenting a goods and services market according to customers' engagement.

The chosen goods and services market is the fuel market of the former 2007 European Capital of Culture - Sibiu, Romania. As data source, end-customers of fuel stations were selected; they represent an external data source (customers who are part of the external environment of the fuel distributing companies), a primary data source (the obtained data is analyzed for the first time by the authors), and also they represent a free data source. A survey was used as a research method and a questionnaire as a research instrument. The sample size was represented by 111 respondents, for which the selection variables were two demographic variables (age [at least 18 years old] and possession of a car) and a behavioral variable (using the car for at least four times per week).

The data were collected in the period beginning with the 1st of March 2011 till the 15th April 2011. Three sets of variables were attracted from each customer and used as input data for RFM scoring model:

i. *Recency of the last purchase* – time period elapsed between the last purchase and the questionnaire completion, denoted by R and measured through a nominal scale with 5 categories: less than 3 days, between 3 and 7 days, between 8 and 11 days, between 12 and 15 days and over 15 days.
ii. *Frequency of purchases* – number of purchases completed by a customer in the last month, denoted by F and measured through a nominal scale with 5 categories: more than 4 times in a month, 4 times in a month, 3 times in a month, 2 times in a month and less than 2 times in a month.
iii. *Monetary value of the last purchase* – the monetary value (in Lei) paid by a customer within his last purchase, denoted by M and measured through a ratio scale.

The k-means clustering procedure uses variables measured only through numeric scales (interval or ratio scales), therefore a transformation of the two nominal variables (R and F) is needed. Numbers (scores) varying from 1 to 5 can be considered alternatives to the categories of the nominal variables. Thus, the mentioned variables are transformed according to the intensity of the customer's engagement, as follows: 1 – most recent transaction (less than 3 days) and 5 – most distant transaction (over 15 days) for *R-recency;* and 1 – lowest frequency (2 times in a month) and 5 – highest frequency (more than 4 times in a month) for *F-frequency*. The M variable was ordered ascending according to its values and the 20th, 40th, 60th and 80th percentiles were computed. The initial values of the M variable were transformed in scores (varying from 1 to 5) based on the computed percentile scores (e.g. if a customer spends an amount of 400 Lei within his last purchase and the 20th percentile is 350 Lei, while the 40th percentile is 430 Lei, then the initial values will be transformed in a score of 2).

After all variables were transformed in scores, the data set was divided in two parts: a training set and a test set. The training set is used to construct the initial clusters based on the RFM variables which are validated or not through the remaining data of the test set. The statistical software tool IBM SPSS V.19 was used to compute three clusters based on the variable scores of the training set.

Table 2. *Cluster centers (and clusters) computed on training set basis*

Initial Cluster Centers	Cluster		
	1	**2**	**3**
Frequency of purchases	2,00	3,23	2,93
Monetary value of the last purchase	2,75	1,86	4,29
Recency of the last purchase	4,38	1,45	1,93

Table 2 contains the initial cluster centers (centroids) computed from the observations of the training set. Three clusters were set as default by the researchers. Based on the values of the centroids, the elements (observations) of the formed clusters can be described as follows:

i. The observations of Cluster 1 are customers which have a low monthly purchase frequency (the value of 2.00 is under the mean value of 2.5). They are middle buyers (the monetary value of their last purchase of 2.75 is near the mean value of 2.5) and their last purchase was far more distant from the time of completing the survey (thus, there is a high probability of purchase in the immediate period).
ii. The observations of Cluster 2 are customers which buy more frequently than the customers of Cluster 1 do, but in a lower value ($3.23 > 2.5 > 2.00$ and $1.86 < 2.5 < 2.75$). These customers are low buyers which have completed their last transaction at a time near the survey period ($1.45 < 2.5$), thus there is a low probability that a purchase will occur in the immediate timeframe.
iii. The observations of Cluster 3 are customers which have a high buying frequency ($2.93 > 2.5$) and the highest spent monetary value for the last purchase (4.29). They are heavy buyers and are the most profitable customers for the company. The company's marketing actions should be oriented towards retaining these heavy buyers.

The distribution of the training set observations, according to their cluster membership, is presented in the table below:

Table 3. *Distribution of the training set observations, according to their cluster membership*

Cluster number	% of observations
1	18.75
2	56.25
3	25.00
Total	**100.00**

The computed centroids were saved and used as starting points in clustering the remaining observations of the test set. A cluster number was attached for all the observations of the initial training set and the test set.

Table 4. *Final clusters centers of the test set observations*

Clustering variables	Cluster		
	1	**2**	**3**
Frequency of purchase	*2.00*	*3.26*	*2.67*
Monetary values of the last purchase	*2.44*	*1.89*	*4.20*
Recency of the last purchase	*4.67*	*1.48*	*2.20*

The final centroids of the three clusters of the test set do not differ so much from the used initial centroids of the formed training set clusters (table 2 and table 4). The clusters of the test set maintain the patterns of the training set clusters, thus each observation (independent of the underlying set – training or test set) is part of one cluster which is consistent through the entire data set. This consistency is understood as non-degeneration (persistency) of the cluster patterns (constant cluster centers) throughout the entire database. The t-Student test was used to prove that this non-degeneration of the cluster patterns is valid throughout the entire data set. The following steps were used:

i. A cluster membership number (1, 2 or 3) was attached to every observation of the entire data set (training set and test set).
ii. The distance from the cluster center was computed for every observation of the entire data set (Euclidean distances were used within a three-dimensional space in which each dimension is represented by one clustering variable – R, F or M).
iii. The data set was split according to the cluster membership and the t-test was applied for the test variable – distance from cluster centers – and a filter (training set/test set) was used as a grouping variable.

The purpose of using this statistical test is to identify if there is any statistical difference between the means of the test variable computed for both the training and the test set. This can be represented by the following symbolic notations:

H_0: $\mu_1 - \mu_2 = 0$
H_1: $\mu_1 - \mu_2 \neq 0$,

where H_0 is the null hypothesis which states that there is no statistical significant difference between μ_1 (the mean of the variable *distances from cluster centers* of the training set observations) and μ_2 (the mean of the variable *distances from cluster centers* of the test set observations) and H_1 the alternative hypothesis which states the opposite.

A validation of H_0 is interpreted as a consistence of the initial formed clusters (based on the training set) throughout the test set. Thus, an observation is a member of one and only one cluster independent of the data that served as a clustering base. In opposition with this aspect, the alternative hypothesis (H_1) presumes that there is insufficient data to confirm the null hypothesis; therefore the initial formed clusters are not consistent throughout the data set.

Table 5. *T-test for the test variable – distances from cluster center – by using the filter - test set/training set - as grouping variable (Cluster 1)*

Levene's Test for Equality of Variances			t-test for Equality of Means						
F		**Sig.**	**T**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**95% Confidence interval of the difference**	
								Lower	Upper
Equal variancesass umed	0.348	0.563	0.526	16	0.606	0.1446396	0.2751497	-0.43	0.72793
Equal variances not assumed			0.526	16	0.606	0.1446396	0.2751497	-0.43	0.72811

The table above contains the results of applying the t-test on the test variable *distances from cluster centers* by using the *membership of the observations* (in the initial training set or test set) as the grouping variable. The Levene's Test for Equality of Variances tests if the variances of the test variable (*distances from cluster centers*) are equal between the two sets of data (training and test set). This is the null hypothesis, which is rejected if the desired significance level is lower than the computed one (Sig. = 0.563). A significance level of Sig. = 0.563 is interpreted as the probability of error if we would reject the null hypothesis; thus the null hypothesis is validated (it can be asserted that there is no statistical difference between the variances of the variable - *distances from cluster centers* computed for the two data sets).

The statistical significant difference between the means of the testing variable (*distances from cluster centers*) considered for the two data sets is tested using the t-Student test. This test denotes a computed t-value of 0.526 and a statistical significance of 0.606. Thus, there is a high probability of error if we reject the null hypotheses (0.606). Therefore we accept it and conclude that there is no statistical significant difference between the means of the testing variables (*distances from cluster centers*) considered for the two data sets (training and test set). Consistency of the initial formed clusters (based on the initial cluster centers of the training set observations) is proven throughout the test set, thus the elements of Cluster 1 of the training set denote similar patterns as the elements of Cluster 1 of the test set do.

Table 6*. T-test for the test variable – distances from cluster center – by using the filter - training set/ test set - as grouping variable (Cluster 2)*

Levene's Test for Equality of Variances			t-test for Equality of Means						
F		**Sig.**	**t**	**df**	**Sig. (2-tail.)**	**Mean Diff.**	**Std. Error Diff.**	**95% Confidence interval of the difference**	
								Lower	**Upper**
Equal variances assumed	0,046	0,831	0,057	54	0,954	0,006	0,113	-0,22	0,233
Equal variances not assumed			0,057	54	0,954	0,006	0,113	-0,22	0,233

The elements of Cluster 2 are more homogenous between the two data sets according to the testing variable: distances from cluster centers. The low value of the t-test (0,057) and the high probability of error if the null hypothesis is rejected (0,954) indicate that there is no statistical significant difference between the mean values (μ_1 and μ_2) of the testing variable considered for the two data sets. A high consistency of the initial formed clusters is proven; the cluster centers computed for the training set observations attract observations of the test set which have strong similar patterns as the first one do. By comparing the results of the t-Student test for Cluster 1 and Cluster 2, a more consistent Cluster 2 is found (0,954 > 0,606) - the elements of Cluster 2 are more homogenous throughout the database than the elements of Cluster 1.

Table 7*. t-test for the test variable – distances from cluster center – by using the filter - training set/ test set - as grouping variable (Cluster 3)*

Levene's Test for Equality of Variances			t-test for Equality of Means						
F		Sig.	T	df	Sig. (2-tail.)	Mean Diff.	Std. Error Diff.	95% Confidence interval of the difference	
								Lower	Upper
Equal variances assumed	2,543	0,121	0,647	31	0,522	0,121	0,187	-0,26	0,503
Equal variances not assumed			0,63	25	0,534	0,121	0,192	-0,27	0,517

Unequal variance of the testing variable throughout the two data sets is denoted by a low Sig. of 0.121 for the Levene's Test. The t-Student test is performed for both equal and unequal variance of the test variable. The results are similar (computed t-values of 0.647 and 0.63 and Sig. of 0.522 and 0.534) and conclude that there is no statistical significant difference between the means of the testing variables considered for the two data sets.

Consistency of the formed clusters was tested throughout the data set by using the cluster centers of the training set as starting centers for the test set. For all the three clusters, consistency was found throughout the data set. Consistency of a cluster can be interpreted as a similarity of patterns or homogeneity of the underlying observations throughout the data set. Based on this idea, the observations of Cluster 1 can be characterized as customers with a low monthly buying frequency (mean frequency of 2.00), with an average monetary value spent within their last purchase (mean monetary values of the last purchase of 2.44), and a high probability of future purchases (their last purchase was far more distant in time - mean recency of last purchase of 4.67).

The Cluster 2 customers are frequent buyers (mean F of 3.26), but light spenders (mean M of 1.89). They have a low probability of completing a transaction in the future (1.48). The third customer group (Cluster 3) consists of high and frequent spenders (mean M of 4.20 and a mean of F of 2.67) which have a medium probability of completing a transaction in the future (mean R of 2.20). These are considered to be the most valuable customers out of the customer database because of their strong engagement with the company's offerings.

3. Conclusion

The purpose of this paper is to define and measure customer engagement. A logical structure was used to define and understand the concept of customer engagement as a part of a broader concept - the value chain (Bruhn, 2010). In this context, customer engagement is defined as the customer's ultimate outcome with causal precedence of satisfaction and trust (Morgan/Hunt 1994 B). An alternative expression for customer commitment is customer engagement which is defined by Van Doorn et al. (2010) as the customer's behavioral manifestation towards a brand or a firm which goes beyond purchase behavior. This behavioral manifestation can be associated with the customers' behavioral effects within the value chain. Based on these two dissimilar points of view, a sensible difference in understanding a customer's commitment can be seized. If commitment is understood as a psychological dimension, then its intensity (formation) is directly linked to other psychological dimensions (such as satisfaction, trust, perceived value) and *customer-related exogenous factors* (company specific – brand reputation). One point of view understands customer commitment through its behavioral manifestations; the other understands customer commitment through its behavioral manifestations. Some of these are visible to the company (purchases), other are not (word of mouth, customer cocreation and complaining behavior). Both of them have an influence over the company's outcome or performance, which can be a direct one (purchase (visible) behavior) or indirect, such as Word-Of-Mouth, loyalty (Bruhn, 2010), participation in the company's activities, customer voice or service improvements (Bijmolt, 2010).

Customer engagement develops and has effects within each stage of the customer lifecycle (customer acquisition stage, customer retention stage, and customer win-back stage (Bruhn, 2010)). For each stage, several models were developed for measuring customer engagement. The authors review two such models specific for the customer selection stage (probability model - Pareto/NBD model and a parametric scoring model - RFM model) and apply one of them (RFM model) on primary data. The purpose of applying the RFM model is to demonstrate that there is no statistical significant variation within the clusters formed on two different data sets (training and test set) if the cluster centroids of the training set are used as initial cluster centroids for the second test set.

Authors' future research will be oriented towards developing and applying persistence models for the measurement of customer engagement. The authors consider such models to be a suitable research instrument for the theoretical framework of the value chain, in general and customer engagement, in particular. The main limitations regarding these models (persistence models, Gupta 2006) can be grouped into two categories. *First,* longitudinal data is necessary in order to apply these kinds of models. Although, there are several tools (especially online tools) which measure customers' dimensions such as satisfaction, trust, etc., on the long-term, these psychological dimensions are not ranked as accurate as transactional data is. *Second,* it is quite difficult to measure and to analyze, in a correct manner, psychological dimensions within transversal marketing researches (one time); conducting longitudinal studies based on cohorts of customers make these tasks more difficult.

4. References

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

Ansari, A. and Mela, C.F., 2003. E-Customization, *Journal of Marketing Research*, 40(2), pp.131-145.

Anderson, E.T. and Simester, D.I., 2004. Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies. *Marketing Science*, 23(1), pp.4-20.

Baesens, B., Viaene, S., Van den Poel, D., Vanthienen, J. and Dedene G., 2002. Bayesian Neural Network Learning for Repeat Purchase Modelling in Direct Marketing. *European Journal of Operational Research*, 138(1), pp.191-211.

Bauer, C.L., 1988. A Direct Mail Customer Purchase Model. *Journal of Direct Marketing*, 2(3), pp.16-25.

Berry, L.L., 1983. Relationship Marketing. *AMA:Emerging Perspectives on Service Marketing*, pp.25-28.

Bijmolt, T.H.A., Leeflang P.S.H., Block F., Eisenbeiss M., Hardie B.G.S., Lemmens A. and Saffert P., 2010. Analytics for Customer Engagement. *Journal of Service Research*, 13(3), pp.341-356.

Bruhn, M., 2009. *Relationship Marketing. Das Management von Kundenbeziehungen. 2.Auflage*. München: Verlag Vahlen.

Bult J.R. and Wansbeek T., 1995. Optimal Selection for Direct Mail. *Marketing Science*, 14(4), pp.378-394.

David Shepard Associates, 1999. *The New Direct Marketing. 3rd Edition*. New York: Irwin Professional Publishing.

Fader, P. S., Hardie B.G.S. and Lee K.K., 2005a. Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science*, 24(2), pp.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), pp.1086-1108.

Forgy, E., 1965. Cluster Analysis of Multivariate Data: Efficiency versus Interpreability of Classifications. *Biometrics*, 21, pp.768.

Grönroos, C., 1990. Relationship Approach to Marketing in Service Contexts. The Marketing and Organizational Behavior Interface. *Journal of Business Research*, 20(1), pp.4-20.

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

Hansonia, B.J. and Wang, P., 1997. Analytical Challenges in Customer Aquisition. *Journal of Direct Marketing*, 11(2), pp. 7-19.

Hughes, A., 2005. *Strategic Database Marketing, 3rd Edition.* New York: McGraw-Hill.

Iacobucci, D. and Ostrom A., 1996. Commercial and interpersonal relationships; Using the structure of interpersonal relationships to understand individual-to-individual, individual-to-firm, and firm-to-firm relationships in commerce. *International Journal of Research in Marketing*, 13, pp.53-72.

Lacey, R., 2007. Relationship Drivers of Customer Commitment. *Journal of Marketing Theory and Practice*, 15(4), 315.

Lewis, M., 2005. Incorporating Strategic Consumer Behavior into Customer Valuation. *Journal of Marketing*, 69(4), pp.230-238.

Malthouse, E.C., 2001. Assesing the Performance of Direct Marketing Scoring Models. *Journal of Interactive Marketing*, 15(1), pp.49-62.

Mobasher, B., Cooley, R. and Srivastava, J., 2000. Automatic Personalization Based on Web Usage Mining. *Communications of the ACM*, 43(8), pp.142-151.

Moe W.W. and Fader P.S., 2004. Dynamic Conversion Behavior at E-Commerce Sites. *Management Science*, 50(3), pp.326-335.

Morgan, R.M. and Hunt S.D., 1994. The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, 58(3), pp.20-38.

Nasraoui, O. and Petenes, C., 2003. *Combining Wen Usage Mining and Fuzzy Inference for Website Personalization*. Proceedings of WebKDD 2003 - KDD Workshop on Web Mining as a Premise to Effective and Intelligent Web Applications. Washington DC, 37.

Oliver, R.L., 1996. *Satisfaction. A Behavioral Perspective on the Consumer*. New York: McGraw-Hill.

Reinartz, W., Krafft, M. and Kumar V., 2005. Balancing Acquisition and Retention Resources to Maximize Customer Profitability. *Journal of Marketing*, 69(1), pp. 63-79.

Sanchez-Garcia, J. and Moliner-Tena M.A., 2007. Relationship Quality of an Establishment and Perceived Value of a Purchase, *Service Industries Journal*, 27(2), pp.151 - 174.

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), pp.1-24.

Sismeiro, C. and Bucklin R.E., 2004. Modeling Purchase Behavior at an E-Commerce Web Site: A Task-Completion Approach. *Journal of Marketing Research*, 41(3), pp.306-323.

Stauss, B., 2008. *Kundenbindung durch Beschwerdemanagement. Handbuch Kundenbindungsmanagement. Strategien und Instrumente für ein erfolgreiches CRM. 6 Auflage.* Wiesbaden, pp.369-396.

Stone, B., 1995. *Successful direct marketing methods*. Lincolnwood, IL: NTC Business Books.

Thomas, J.S., Blattberg, J.R. and Fox, E.J., 2004. *Recapturing Lost Customers*. Journal of Marketing Research, 41(1), pp.31-45.

Ting, I.-H., Kimble C. and Kudenko D, 2005. UBB Mining: *Finding Unexpected Browsing Behavior in Clickstream Data to improve a Web Site's Design*. Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence, France, pp.179-185.

Van Doorn, J., Lemon K., Mittal V., Nass S., Pick D., Pimer P. and Verhoef P., 2010. Customer Engagement Behavior: Theoretical Foundations and Research Directions. *Journal of Service Research*, 13(3), pp.253-266.

Verhoef, P.C. and Donkers, B., 2005. The Effect of Aquisition Channels on Customer Loyalty and Cross-Buying. *Journal of Interactive Marketing*, 19(2), pp.31-43.

Vroomen, B., Donkers B., Verhoef P.C. and Franses P.H., 2005. Selecting Profitable Customers for Complex Services on the Internet. *Journal of Service Research*, 8(1), pp.37-47.

Zeithaml, V.A., Parasuraman A. and Berry L.L., 1992. *Quality Service*. Frankfurt am Main/New York.