J. Ind. Eng. Int., 7 (12), 19-31, Winter 2011
ISSN: 1735-5702

Customer lifetime value model in an online toy store

B. Nikkhahan [1]; A. Habibi Badrabadi [2]; M.J. Tarokh [3*]

[1.2] *Postgraduate student, IT Group, Dep. of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran*

[3]*Associate Professor, IT Group, Dep. of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran*

Received: 9 April 2009; Revised: 1 June 2009; Accepted: 9 June 2009

Abstract: Business all around the world uses different approaches to know their customers, segment them and formulate suitable strategies for them. One of these approaches is calculating the value of each customer for the company. In this paper by calculating Customer Lifetime Value (CLV) for individual customers of an online toy store named Alakdolak, three customer segments are extracted. The level of profitability for customers is identified, and finally suitable marketing strategies for each segment are developed. The results indicate that the company should increase its low price products and develop special programs for those that buy high price products and have high loyalty. Logistic regression as a data mining technique is used to present the customer defection and future purchase probability models and for each model, verification and validation is done.

Keywords: *Customer lifetime value; Data mining; Online store*

1. Introduction

In the past two decades, the firms tended to focus on either cost management or revenue growth. When a firm adopts one of these approaches it looses out on the other (Rust *et al.,* 2004). Assessing the value of individual customers and employing customer level strategies based on customers' worth to the firm, leads to effective allocation of resources and efforts across profitable customers. Many customer oriented firms realize that the customers are valued more than the profit they bring in every transaction. Customers' value has to be based on their contribution to the firm across the duration of their relationship with the firm. In simple terms, the value of a customer is the value the customer brings to the firm over his/her lifetime (Kumar 2006).

Valuing customers is a central issue of any commercial activity. The value of an individual customer is important for the detection of the most valuable ones, which deserve to be closely followed, and for the detection of the less valuable ones, to which the company should pay less attention. At the aggregated level, a marketing campaign targeting a group of customers can be budgeted more efficiently when the value of this group is known. Customers are an important asset, and as such, have to be precisely valuated.

Customer Lifetime Value (CLV) plays a major role in customer valuation, in particular Churn Analysis and Retention Campaign Management. In the context of Churn Analysis, the CLV of a customer or a segment is important complementary information to their churn probability, as it gives a sense of how much is really being lost due to churn and how much effort should be concentrated on this segment (Rosset *et al*., 2003). Several authors have developed models that focus on the allocation of scarce marketing resources based on CLV (e.g., Reinartz *et al.,* 2005; Rust *et al.,* 2004; Venkatesan and Kumar, 2004). These approaches use CLV to develop a rank order of customers and recommend devoting more resources to customers with higher ranks. Some authors go so far as to state that customers with low ranks (especially when these show negative CLV) should be abandoned completely to increase overall profitability of the customer base (Zeithaml *et al.*, 2004).

Many definitions exist for store image and store attributes. Verhagen and Van Dolen (2009) believe that their essence is a total impression of tangible or functional factors (merchandise selection, prices ranges and store layout) and intangible or psychological factors (such as perceived manner of the sales staff, service level, and reputation). They used a multi-dimensional construct consisting of: service, merchandise, atmosphere, and layout, dealing therefore with overall impressions of both bricks and clicks but adding customer's perception of ease of online navigation. Van Der Heijden and Verhagen (2004) presented multiple-item measurements for components of a store image: online store usefulness, enjoyment, ease of use, store style, familiarity, trustworthiness, and settlement performance. The components were regressed on attitudes and intentions

*Corresponding Author Email: mjtarokh@kntu.ac.ir
Tel.: +98 9123844762

towards purchasing at the online store, revealing significant, direct influences from usefulness, enjoyment, trustworthiness and settlement performance. Da Silva and Syed Alwi (2008) found that factors such as ease of use, "personalization", security and customer care are significant in determining the corporate brand image of the online e-tailer.

2. Fundamentals of CLV modeling

There are some definitions for CLV in Table1. CLV is generally defined as the present value of all future profits obtained from a customer over his or her life of relationship with a firm. A simple mathematical model of CLV for a customer (omitting customer subscript) is (Tarokh *et al*., 2006; Gupta *et al*., 2004; Reinartz and Kumar, 2003; Sheth *et al*., 2002):

$$CLV = \sum_{i=1}^{n} \frac{R_i - C_i}{(1+d)^i} \quad (1)$$

where i is period of cash flow from customer transactions; R_i the revenue from customer in period i; C_i the total cost of generating the revenue R_i in period i; d interest rate; and n the total number of periods of projected life of customer under consideration. Therefore the numerator is the net profit that has been obtained at each period while the denominator transforms the profit into the current value. This model is the most basic model that ignores the function of sales and costs. Expanding the model with data mining techniques can consider future customers' purchase behavior.

Researchers either build separate models for customer acquisition, retention, and margin or sometimes combine two of these components. For example, Thomas (2001) and Reinartz *et al.* (2005) simultaneously captured customer acquisition and retention. Fader *et al.* (2005) captured recency and frequency in one model and built a separate model for monetary value. However, the approaches for modeling these components or CLV differ across researchers. Gupta *et al.* (2006) described six modeling approaches:

- RFM models,
- Probability models,
- Econometric models,
- Persistence models,
- Computer science models,
- Diffusion / Growth models.

RFM models create "cells" or groups of customers based on three variables: Recency, Frequency, and Monetary value of their prior purchases. Several recent studies have compared CLV models (discussed later) with RFM models and found CLV models to be superior. Reinartz and Kumar (2003) used a catalog retailer's data of almost 12,000 customers over 3 years to compare CLV and RFM models. They found that the revenue from the top 30% of customers based on the CLV model was 33% higher than the top 30% selected based on the RFM model.

Venkatesan and Kumar (2004) also compared several competing models for Customer Selection. Using data on almost 2,000 customers from a business-to-business (B2B) manufacturer, they found that the profit generated from the top 5% customers as selected by the CLV model was 10% to 50% higher than the profit generated from the top 5% customers from other models (e.g., RFM, past value, etc.).

A probability model is a representation of the world in which observed behavior is viewed as the realization of an underlying stochastic process governed by latent (unobserved) behavioral characteristics, which in turn vary across individuals. Schmittlein and Peterson (1994), Colombo and Jiang (1999), Reinartz and Kumar (2000, 2003), Fader *et al.* (2004) have all proposed probability models to compute CLV.

Table 1: LTV definitions.

The NPV of all future contributions to overhead and profit.	Roberts and Berger (1989)
NPV of a future stream of contributions to overheads and profits.	Jackson (1994)
The NPV of all future contributions to profit and overhead expected from the customer.	Courtheoux (1995)
The NPV of the stream of contributions to profit that results from customer transactions and contacts.	Pearson (1996)
Expected profit from customers, exclusive of cost related to customer management.	Blattberg and deighton (1996)
The total discounted net profit that a customer generates during her life on the house list.	Bitran and Mondschein (1996)
The present value of all future profits generated from a customer.	Gupta and Lehmann (2003)
Sum of cumulated cash flows - discounted using the Weighted Average Cost of Capital (WACC) - of a customer over his or her entire lifetime with the company.	Kumar (2006)
Estimated monetary value that the client will bring to the firm during the entire lifespan of his/her commercial relationship with the company, discounted to today's value.	Yamamoto Sublaban and Aranha (2008)

Econometric models deal with customer acquisition, retention, and expansion (cross-selling or margin) and then combine them to estimate CLV.

Like econometric models of CLV, persistence models focus on modeling the behavior of its components, that is, acquisition, retention, and cross-selling. When sufficiently long-time series are available, it is possible to treat these components as part of a dynamic system. To date, this approach, known as persistence modeling, has been used in a CLV context to study the impact of advertising, discounting, and product quality on customer equity (Yoo and Hanssens 2005) and to examine differences in CLV resulting from different customer acquisition methods (Villanueva *et al.*, 2006).

The vast computer science literature in data mining, machine learning, and nonparametric statistics has generated many approaches that emphasize predictive ability. These include projection-pursuit models; neural network models; decision tree models; spline-based models such as Generalized Additive Models (GAM), Multivariate Adaptive Regression Splines (MARS), Classification and Regression Trees (CART); and Support Vector Machines (SVM).

CLV is the long-run profitability of an individual customer. This is useful for Customer Selection, Campaign Management, Customer Segmentation, and Customer Targeting (Kumar 2006b). Whereas these are critical from an operational perspective, CLV should be aggregated to arrive at a strategic metric that can be useful for senior managers. With this in mind, several researchers have suggested that we focus on Customer Equity (CE), which is defined as the CLV of current and future customers (Blattberg *et al.*, 2001; Gupta and Lehmann 2005; Rust *et al.*, 2004).

3. Improved CLV model

In spite of simple formulation (Equation 1), researchers have used different variations in modeling and estimating CLV. Some researchers have used an arbitrary time horizon or expected customer lifetime (Reinartz and Kumar 2000; Thomas 2001), whereas others have used an infinite time horizon (e.g., Fader *et al.,* 2005; Gupta *et al.,* 2004). Gupta and Lehmann (2005) showed that using an expected Customer Lifetime generally overestimates CLV, sometimes quite substantially. Calculating LTV with the basic model has some deficiencies:

- It cannot consider customer churn probability,
- It cannot consider factors as cress-selling opportunity,
- The possibility of up-selling cannot be calculated.

Hwang (2003) suggested a complimentary model (Equation 2):

$$LTV_i = \sum_{t_i=1}^{N_i} \pi_P(t_i)(1+d)^{N_i - t_i} + \sum_{t_i=N_i+1}^{N_i+1+E(i)} \frac{\pi_f(t_i) + B(t_i)}{(1+d)^{t_i - N_i}} \qquad (2)$$

where:

T_i Service period index of customer *i*.

N_i Total service period of customer *i*.

d Interest rate.

$E(i)$ Expected service period of customer.

$\pi_p(t_i)$ Past profit contribution of customer *i* at period t_i.

$\pi_f(t_i)$ Future profit contribution of customer *i* at period t_i.

$B(t_i)$ Potential benefits from customer *i* at period t_i.

$\pi_p(t_i)$ Can be calculated from Equation (1).

$\pi_f(t_i)$ Is the prediction of $\pi_p(t_i)$ in the future and can be calculated using regression techniques.

In order to calculate $E(i)$ we should describe the relation between expected service period of customers and customer defection (churn) rate.

Let *y* be the number of service time required until a customer service is stopped and let P_{Churn} be the churn probability of a customer. Then, the probability of mass function of service period is given by:

$$P(n) = P(y = n) = P_{Churn} \times (1 - P_{Churn})^{n-1}, n = 1,2,3,...$$
$$\therefore y \sim \text{Geo}(P_{Churn}) \Rightarrow E[i] = \frac{1}{P_{Churn}} \qquad (3)$$

In order to calculate $B(t_i)$, we should use Equation (4) as follows:

$$B(t_i) = \sum_{j=1}^{k} prob_{ij} \times prof_{ij} \qquad k = 1,2,3,$$

= Potential benefits of customer in period i (4)

where $prob_{ij}$ is the probability that customer i would use the service j among k-potential services. $prof_{ij}$ means the profit that a company can get from customers i who uses optional service j. it is available by substituting the cost of each optional service from the charge of each optional service. However, we should use data mining techniques in order to calculate $prob_{ij}$.

4. Case study: Online toy store

In this study, data were gathered from an online store. Alakdolak is an online store that sells toys. Children are its customers but most of the time other people buy from the store for their relatives. Children are indirect customers, but we need their age to predict their relatives' behavior in online store. So we consider children as customers. Our data includes two types of data. One of them is personal information like age, sex, relation of the buyer to the customer (like son, daughter). Another one is the shopping history of the customers like the name of the product, its cost and the date of shopping.

The data are related to 182 customers and their 1935 purchasing records in the last five months of 2008. First we need to calculate the probability of customer churn and the probability of buying the different types of the products by the customers. Then we can segment the customers in order to suggest the marketing strategies for the online store. Simply from the raw data, the following information can be concluded:

- Number of the products that each customer has bought,
- Sum of the customer shopping cost from the site,
- Frequency of customer shopping,
- Date of customer's last shopping.

In this study the researchers want to calculate the probability of customer churn.

So they must have a definition for a defected customer. For this purpose they set up an interview with the managers of the store. They believed that customers of this online store are children and they have limited period of using toys and most of the parents buy toys for their children in their childhood. So if they didn't buy any products from the shop for three months, we can assume that they are buying from other stores.

We can categorize the products into three categories: Low, Medium and High Price. Now we want to calculate the probability of buying these products by customers. Based on the case the researchers present four models: A model for calculating the customer defection probability, and 3 models for calculating cross-selling probability.

5. Modeling and data mining

We dealt with two types of variables for prediction; type one is the probability of customer churn (Buying the products) and type two is the possibility of purchasing each product type by the customer. Logistic regression was used as a data mining technique. A table of 182 customers was used with 11 Data fields. The churn behavior of customers recorded as a binary variable. Equation (5) represents the logistic regression (Menard and Scot 1995). Where β_0 is the constant value and β_i is model parameters.

$$P(y=1) = \frac{1}{1+\exp^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+...)}} \qquad (5)$$

5.1. Calculating customer defection probability

The researchers used to SPSS software to establish logistic regression model. Customer churn data were chosen as the response variable and backward stepwise method with Wald statistics was implemented. In the first step, all variables entered into the model and then through hypothesis test H_0: $\beta_i=0$ some of them eliminated. In the first model, B is independent variable coefficient. Using Microsoft Excel, the probability of customer churn was calculated for each customer.

5.2. Calculating cross-selling probability

All the 3 types of products were considered for cross-selling opportunity. It consists of 3 models (one model for each product type) with SPSS software as follows:

Classification Table[a]

	Observed		Predicted		
			Churn		Percentage
			0	1	Correct
Step 1	churn	0	153	1	99.4
		1	2	26	92.9
	Overall Percentage				98.4
Step 2	Churn	0	153	1	99.4
		1	2	26	92.9
	Overall Percentage				98.4
Step 3	churn	0	153	1	99.4
		1	2	26	92.9
	Overall Percentage				98.4
Step 4	churn	0	152	2	98.7
		1	4	24	85.7
	Overall Percentage				96.7
Step 5	churn	0	153	1	99.4
		1	4	24	85.7
	Overall Percentage				97.3

a. The cut Value is .500

Variables in the Equation

		B	SE	Wald	df	Sig	Exp(B)
Step 1[a]	Sale date max	-.315	.096	10.802	1	.001	.730
	Sale count	-2.402	2.829	.721	1	.396	.090
	Sex	1.663	1.452	1.311	1	.252	5.276
	RelationId	1.814	.662	7.509	1	.006	6.137
	Age	.003	.008	.167	1	.683	1.003
	Saler product cpimt	.177	.328	.292	1	.589	1.194
	Sale product cost sum	.000	.000	1.184	1	.276	1.000
Step 2[a]	Sale date max	-.311	.094	10.877	1	.001	.733
	Sale count	-2.594	2.787	.867	1	.352	.075
	Sex	1.595	1.440	1.226	1	.268	4.927
	RelationId	1.843	.671	7.542	1	.006	6.315
	Sale product count	.211	.323	.427	1	.513	1.235
	Sale product cost sum	.000	.000	1.299	1	.254	1.000
Step 3[a]	Sale date max	-.314	.093	11.297	1	.001	.731
	Sale count	-2.381	2.727	.762	1	.383	.092
	Sex	1.317	1.317	1.000	1	.317	3.734
	RelationId	1.868	.679	7.568	1	.006	6.475
	Sale product cost sum	.000	.000	1.625	1	.202	1.000
Step 4[a]	Sale date max	-.306	.088	12.006	1	.001	.737
	Sale count	1.243	1.285	.937	1	.333	3.467
	Sex	1.498	.468	10.240	1	.001	4.472
	Sale product cost sum	.000	.000	5.971	1	.015	1.000
Step 5[a]	Sale date max	-.288	.081	12.639	1	.000	.750
	RelationId	1.589	.475	11.187	1	.001	4.900
	Sale product cost sum	.000	.000	6.572	1	.010	1.000

a. Variable(s) entered on Step1: Sale date max, Sale count, Sex, RelationId, Age, Sale Product Count, Sale Product Cost Sum.

Model 1: Prediction of customer churn.

Classification Table[a]

	Observed		Predicted		
			Low Price		Percentage
			0	1	Correct
Step 1	Low Price	0	88	12	88.0
		1	7	75	91.5
	Overall Percentage				89.6
Step 2	Low Price	0	90	10	90.0
		1	7	75	91.5
	Overall Percentage				90.7
Step 3	Low Price	0	89	11	89.0
		1	8	74	90.2
	Overall Percentage				89.6
Step 4	Low Price	0	85	15	85.0
		1	6	76	92.7
	Overall Percentage				88.5
Step 5	Low Price	0	85	15	85.0
		1	7	75	91.5
	Overall Percentage				87.9

a. The cut Value is .500

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)
Step 1[a]	Sex	.699	.541	1.668	1	.197	2.012
	RelationId	.052	.156	.111	1	.739	1.053
	Age	.017	.011	2.347	1	.126	1.017
	Saler product count	-.998	.169	34.717	1	.000	.369
	Sale count	2.932	1.152	6.476	1	.011	18.766
	Sale Date Max	.006	.006	1.044	1	.307	1.007
Step 2[a]	Sex	.712	.540	1.741	1	.187	2.039
	Age	.017	.011	2.402	1	.121	1.017
	Sale Product Count	-1.005	.169	35.211	1	.000	.366
	Sale Count	3.203	.807	15.752	1	.000	24.609
	Sale Date Max	.007	.006	1.108	1	.293	1.007
Step 3[a]	Sex	.731	.537	1.853	1	.173	2.077
	Age	.017	.010	2.721	1	.099	1.017
	Sale Product Count	-1.000	.167	35.864	1	.000	.368
	Sale count	3.594	.725	24.587	1	.000	36.380
Step 4[a]	Age	.018	.010	3.083	1	.079	1.018
	Sale Product Count	-1.002	.165	37.038	1	.000	.367
	Sale count	4.010	.678	34.943	1	.000	55.129
Step 5[a]	Sale Product Count	-.964	.157	37.682	1	.000	.381
	Sale count	4.246	.675	39.516	1	.000	69.802

a. Variable(s) entered on Step1: Sex, RelationId, Age, Sale Product Count, Sale count, sale date Max.

Model 2: Prediction of buying the low price products.

Classification Table[a]

	Observed		Predicted		
			Medium Price		Percentage
			0	1	Correct
Step 1	Medium Price	0	89	19	82.4
		1	51	23	31.1
	Overall Percentage				61.5
Step 2	Medium Price	0	90	18	83.3
		1	50	24	32.4
	Overall Percentage				62.6
Step 3	Medium Price	0	90	18	83.3
		1	53	21	28.4
	Overall Percentage				61.0
Step 4	Medium Price	0	93	15	86.1
		1	52	22	29.7
	Overall Percentage				63.2
Step 5	Medium Price	0	86	22	79.6
		1	54	20	27.0
	Overall Percentage				58.2
Step 6	Medium Price	0	108	0	100.0
		1	74	0	.0
	Overall Percentage				59.3

a.The cut Value is .500

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)
Step 1[a]	Sex	-.428	.303	1.997	1	.158	.652
	RelationId	.156	.075	4.377	1	.036	1.169
	Age	-.004	.004	.847	1	.357	.996
	Saler product countt	.003	.009	.095	1	.758	1.003
	Sale count	-.369	.339	1.185	1	.276	.691
	Sale Date Max	-.006	.004	2.757	1	.097	.994
Step 2[a]	Sex	-.426	.303	1.975	1	.160	.653
	RelationId	.153	.073	4.322	1	.038	1.165
	Age	-.004	.004	.881	1	.348	.996
	Sale Count	-.329	.310	1.125	1	.289	.720
	Sale Date Max	-.006	.004	2.755	1	.097	.994
Step 3[a]	Sex	-.447	.301	2.201	1	.138	.639
	RelationId	.148	.074	4.046	1	.044	1.160
	Sale count	-.389	.314	1.532	1	.216	.678
	Sale Date Max	-.006	.003	2.685	1	.101	.994
Step 4[a]	Sex	-.474	.299	2.516	1	.113	.622
	RelationId	.094	.057	2.687	1	.101	1.098
	Sale Date Max	-.008	.003	5.411	1	.020	.993
Step 5[a]	RelationId	.052	.050	1.100	1	.294	1.054
	Sale Date Max	-.008	.003	6.203	1	.013	.992
Step 6[a]	Sale Date Max	-.005	.002	9.317	1	.002	.995

a.Variable(s) entered on Step1: Sex, RelationId, Age, Sale Product Count, Sale count, sale date Max.

Model 3: Prediction of buying the medium price products.

Classification Table[a]

Observed			Predicted		
			High Price		Percentage Correct
			0	1	
Step 1	High Price	0	153	3	98.1
		1	3	23	88.5
	Overall Percentage				96.7
Step 2	High Price	0	152	4	97.4
		1	3	23	88.5
	Overall Percentage				96.2
Step 3	High Price	0	152	4	97.4
		1	3	23	88.5
	Overall Percentage				96.2
Step 4	High Price	0	153	3	98.1
		1	3	23	88.5
	Overall Percentage				96.7
Step 5	High Price	0	151	5	96.8
		1	2	24	92.3
	Overall Percentage				96.2

a.The cut Value is .500

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)
Step 1[a]	Sex	-.672	.888	.574	1	.449	.511
	Relationld	-1.448	.344	17.671	1	.000	.235
	Age	.002	.006	.151	1	.698	1.002
	Sale product count	.323	.081	16.034	1	.000	1.382
	Sale count	.890	.863	1.063	1	.303	2.436
	Sale Date Max	-.003	.009	.131	1	.717	.997
Step 2[a]	Sex	-.819	.800	1.050	1	.306	.441
	Relationld	-1.446	.338	18.295	1	.000	.236
	Age	.002	.006	.197	1	.657	1.002
	Sale product count	.318	.077	16.875	1	.000	1.374
	Sale count	.819	.842	.946	1	.331	2.268
Step 3[a]	Sex	-.813	.805	1.020	1	.313	.443
	Relationld	-1.437	.331	18.887	1	.000	.238
	Sale product count	.313	.075	17.347	1	.000	1.367
	Sale count	.916	.801	1.307	1	.253	2.499
Step 4[a]	Relationld	-1.480	.332	19.835	1	.000	.228
	Sale product count	.327	.075	19.154	1	.000	1.387
	Sale count	.592	.699	.716	1	.397	1.09
Step 5[a]	Relationld	-1.359	.283	22.991	1	.000	.257
	Sale product count	.343	.072	22.576	1	.000	1.409

a.Variable(s) entered on Step1: Sex, Relationld, Age, Sale Product Count, Sale count, sale date Max.

Model 4: Prediction of buying the high price products.

6. Model verification

There are several statistics which can be used for comparing alternative models or evaluating the performance of a single model (Menard and Scott, 1995).

6.1. Model Chi-square

In logistic regression, the statistic is analogous to the F-test in linear regression (Menard and Scott, 1995). By multiplying the log-likelihood by -2 it approximates a distribution (Menard and Scott, 1995).

$$G_M = D_0 - D_M \tag{6}$$

where D_0 is initial log likelihood and D_M is final log likelihood. D_0 is analogous to the total sum of square (SST) in linear regression analysis and D_M is analogous to the Error sum of square (SSE).

$$G_M \sim X^2_{0.99,k}$$

where k is degree of freedom and is equal to the number of remaining variables in the final step of modeling. For example for model 1, $G_M = 225.744$ and because $G_M > X^2_{0.99,2} = 9.210$ the null hypothesis is rejected and model significance is verified.

7. Model accuracy

SPSS creates classification table for each step in the logistic regression model. Each classification table summarizes the observed and predicted values to interpolate predictive efficiency/ accuracy for each step in the model. The bigger the percentage correct the better the model. The classification table is presented following each Model 1-4.

7.1. Pseudo- R^2 statistics

It is the proportion of the variance in the dependent variable which is explained by the variance in the independent variables. There are several Pseudo- R^2 statistics. Analogous to $R^2 = \frac{SSR}{SST}$ for linear regression is Pseudo- R^2 in logistic regression (Menard and Scott, 1995). Pseudo- R^2 indicates how much the inclusion of the independent variable in the model reduces the badness of the fit. It varies between 0 (independent variable is useless in the prediction of the dependent variable) and 1 (Independent variable in the model predicts the dependent variable perfectly) and is calculated by dividing the model chi-square (G_M) by the initial log-likelihood function (Menard and Scott, 1995).

$$\text{Pseudo - R}^2 = \frac{G_M}{D_0} = \frac{G_M}{G_M + D_M} \tag{7}$$

Pseudo-R^2 is calculated for each Model 1-4 and variables eliminated by SPSS at each step are tested. For example conclusion for Model 4 (High Price) is shown in Table 5.

Table 2: R2 changes for churn model (Model 1).

Step	-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
1	24.109[a]	.715	.953
2	24.281[a]	.714	.952
3	24.730[a]	.714	.951
4	25.563[a]	.712	.950
5	26.562[b]	.711	.948

a. Estimation terminated at iteration number 11 because parameter estimates changed by less than .001.
b. Estimation terminated at iteration number 10 because parameter estimates changed by less than .001.

Table 3: R2 changes for low price model (Model 2).

Step	-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
1	91.380[a]	.587	.783
2	91.493[a]	.587	.782
3	92.632[a]	.584	.779
4	94.507[a]	.580	.773
5	97.723[a]	.572	.763

a. Estimation terminated at iteration number 9 because parameter estimates changed by less than .001.

Table 4: R2 changes for medium price model (Model 3).

Step	-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
1	235.959[a]	.086	.115
2	236.053[a]	.085	.114
3	237.012[a]	.051	.107
4	238.890[b]	.071	.095
5	241.419[b]	.058	.077
6	242.526[c]	.052	.070

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.
b. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.
c. Estimation terminated at iteration number 2 because parameter estimates changed by less than .001.

Table 5: R2 changes for high price model (Model 4).

Step	-2 Log Likelihood	Cox & Snell R Square	Nagelkerke R Square
1	39.393[a]	.690	.919
2	39.525[a]	.689	.919
3	39.719[a]	.689	.919
4	40.773[a]	.687	.916
5	41.514[a]	.686	.915

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

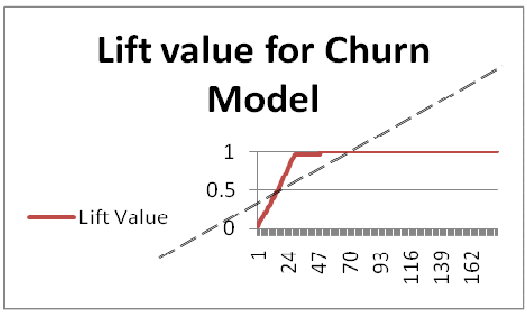

Figure 1: Lift chart for churn model (Model 1).

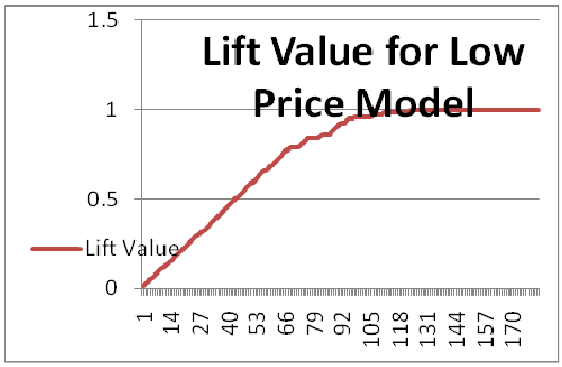

Figure 2: Lift chart for low price model (Model 2).

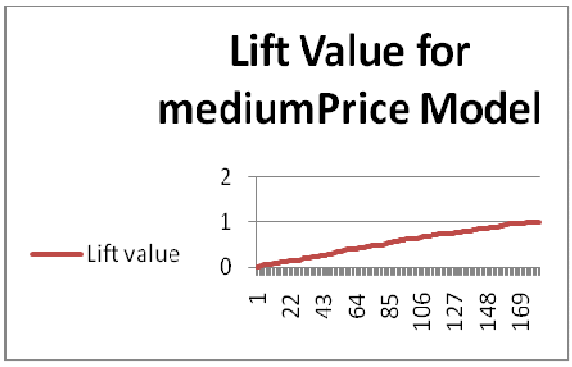

Figure 3: Lift chart for medium price model (Model 3).

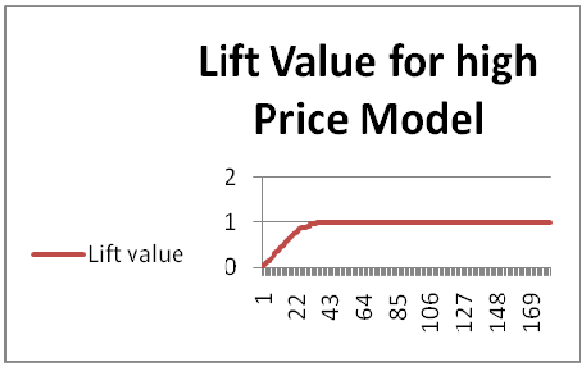

Figure 4: Lift chart for high price model (Model 4).

8. Model validation

Lift chart is one of performance test techniques (Hwang, 2003). The lift value is the ratio of the number of customers who actually left the company within the certain section to the total number of customers who left when we divided sections by sorting target customers in ascending order of high churn rate. The lift chart as proposed above generally shows the cumulative percentage of deviated customers. The researchers conclude that the reliability of the model is good when the graph is skewed to the left. In some notes the base line is mentioned as the measure of reliability (Hwang, 2003). As showed in Figures 1 to 4, "churn model", "low price model", and "high price model" have better results unlike "medium price model". So we can conclude that the model of buying medium price products does not reflect the actual behavior of customers truly.

9. Discussion

Through logistic regression modeling, we can calculate Equation (2) and consider three types of value for customers. As customer life cycle has been considered for five months and we have traced customer purchase behavior in five months, the present value of customers concluded without discount rate. The present value of customers for the online store is sum of their shopping cost.

Customer future value is calculated with Equation (4), considering logistic regression Models 2, 3, and 4. We consider customer loyalty as another measure to segment the customers. We have calculated the customer churn probability through Model 1. So we can define:

Customer Loyalty = 1 – Churn probability

To distinguish between customer segments we used a 3D diagram showed in Table 4.

As shown in Figures 5 and 6, we can categorize the customers of the online store in three main segments:

1. Those that their current value, future value and loyalty are below the average,
2. Those that their current value and future value are below the average, and their loyalty is above the average,
3. Those that their current value, future value and loyalty are above the average.

First customer segment consists of those that their frequency and cost of buying are low. Most of them shopped from the site just one time and they never came back.

Customers in the second segment, generate most of the customer value of the store. Most of them bought cheap products, but their loyalty is high.

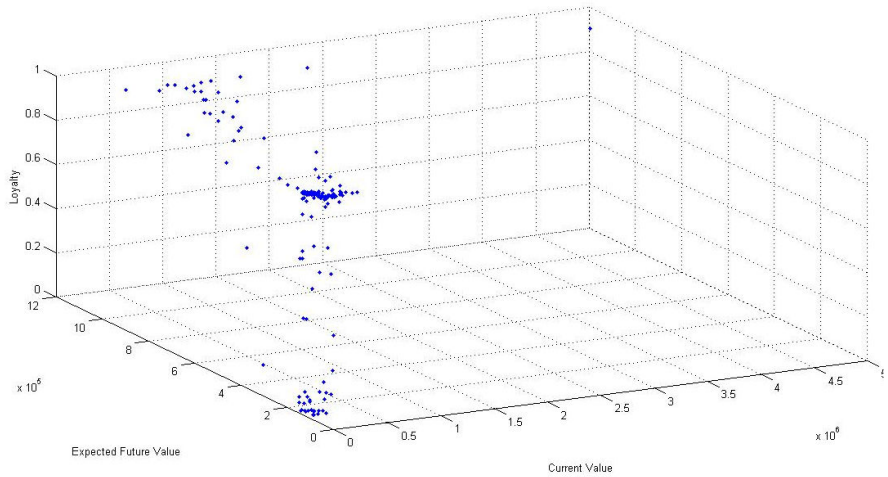

Figure 5: Customer segments.

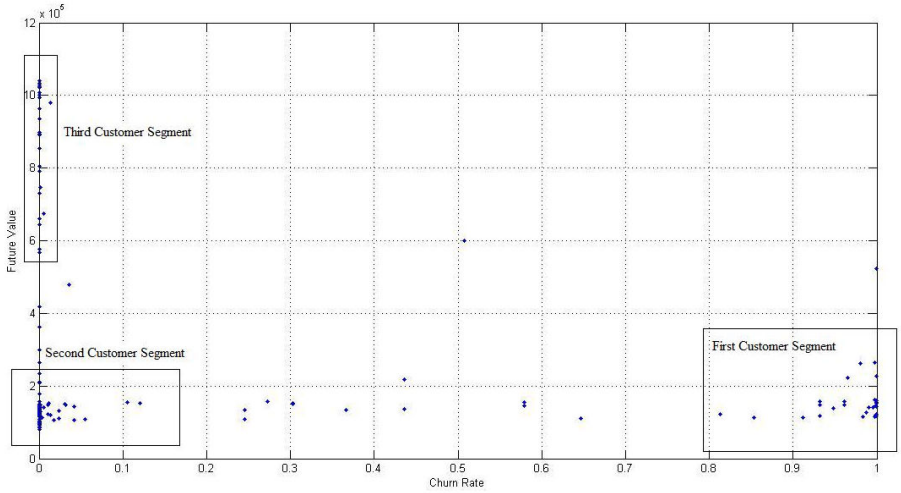

Figure 6: Future expected value/customer churn rate.

Table 6: Segmentation of customers based on Customer Lifetime Profits and relationship duration (Reinartz *et al*., 2002).

Customer Lifetime Profits	Low Relationship Duration	High Relationship Duration
High	**BUTTERFLIES** • Good fit between company's offerings and customers' needs • High Profit potential • Action o Aim for transactional satisfaction, not attitudinal loyalty o Maximize profits from these accounts as long as they are active o Stop investing once inflection point is reached	**TRUE FRIENDS** • Excellent fit between company's offerings and customers' needs • Highest profit potential • Action o Consistent intermittently spaced communication o Achieve attitudinal and behavioral loyalty o Invest to nurture/defend/retain
Low	**STRANGERS** • Little fit between company's offerings and customers' needs • Lowest profit potential • Action o Make no investment in these relationships o Make profit on every transaction	**BARNACLES** • Limited fit between company's offerings and customers' needs • Low profit potential • Action: o Measure size and share of wallet o If share-of-wallet is low, focus on specific up and cross selling o If size of wallet is small, impose strict cost controls

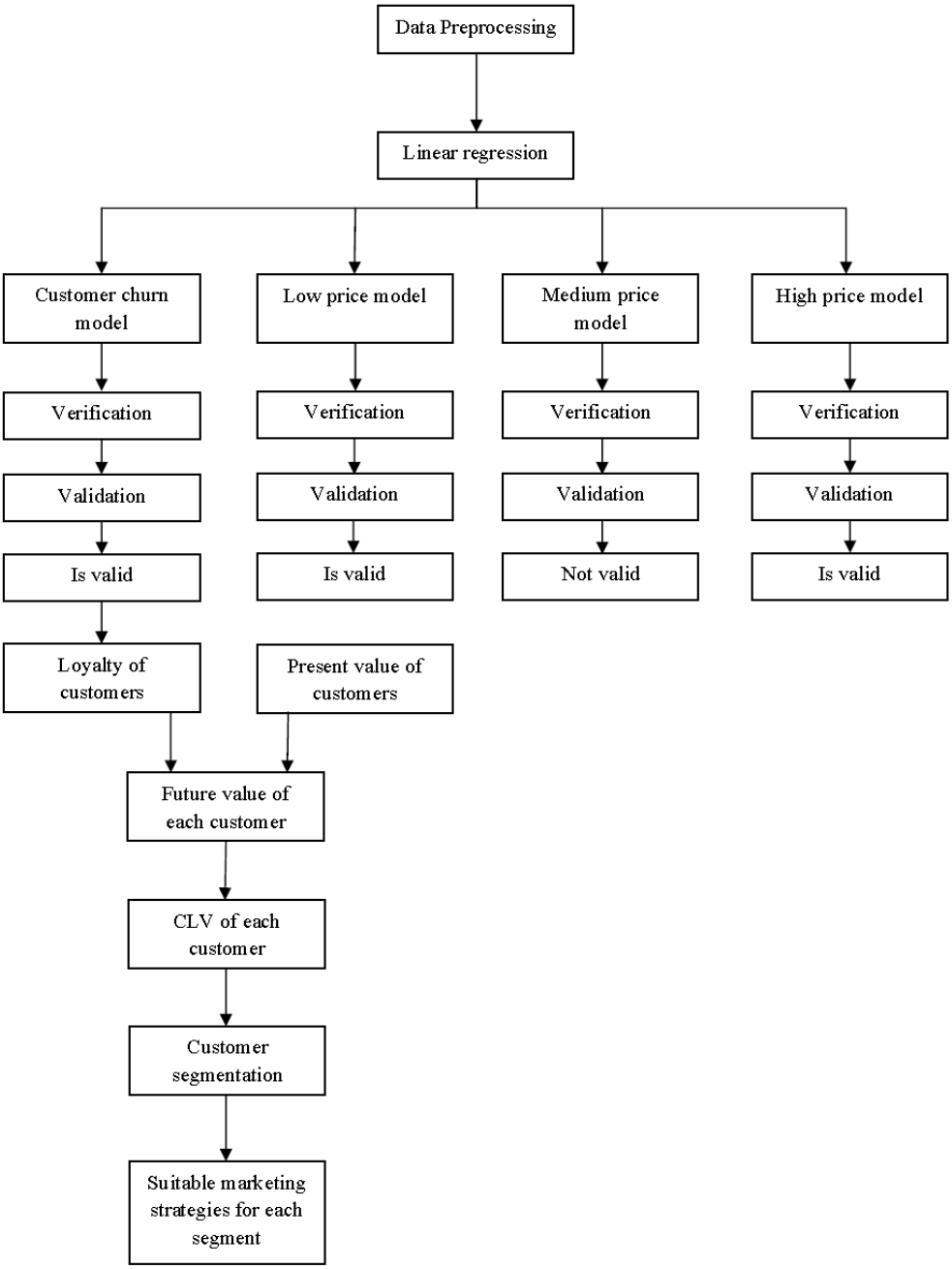

Figure 7: Overview of the study.

We expect that they will buy products from the store again in the future. This customer segment approximately consists of half of the store's customers. Third customer segment are store's special customers. This segment and the first segment are approximately equal in amount.

They are generating more value than the others for the store; we can expect them to generate this value in the future, because they have high future value. In spite of this, because of their loyalty, the store can sell more products to this customer segment. Reinartz *et al*. (2002) presented a segmentation of customers based on Customer Lifetime Profits and relationship duration. This model is used to develop suitable marketing strategies for the store. As shown in Table 6 customers are segmented by their value for the firm and their relationship duration. We can locate our segments on this table. We categorized the customers by their loyalty and their value for the store. So our first segment includes strangers, second segment includes barnacles and third segment includes true friends, and we don't have any butterflies.

The company must increase low price products because most of the store's profits belong to the barnacles. Each one of them generates small value, but they have high loyalty and they are more than 50 percent of the customers. So totally, it is

the most valuable customer segment for the company. The true friends are special customers of the company. They buy expensive products and they have high loyalty, but they are a small portion of the customers. As showed in Table 6, the company should develop special programs to have better interaction with them. Most of the strangers buy just one time and they generate small values. The company must increase the number of true friends and barnacles; as a result it leads to decrease the number of strangers.

An overview of all tasks that are performed in this study is presented in Figure 7. At the first step, raw data must become ready for building the models. In the preprocessing of data, many tasks are done. Some concepts are defined such as customer churn, low, medium and high price products. For example a defected customer is one that didn't buy anything from the store since three month ago. Another task in preprocessing is converting raw data to useful date. For example deriving number of purchases of each customer, sum of its purchases, number of days from the last purchase and so on. After making the data ready, they imported to SPSS software and with linear regression, four models were built. Then we verified and validated the models. In validation process we found out that medium price model is not valid, but the others are valid.

In order to calculate the future value of each customer, we need his or her present value and probability of defection. So for each customer, future value was calculated. CLV of customer calculated by their present and future values. Then the customers segmented by their loyalty, present and future values. Finally based on the segmentation suitable strategies are generated to dealing with different customer segments.

10. Conclusion

Businesses want to allocate their scarce resources to their customers effectively. They want to evaluate and categorize different segments of customers, and then formulate suitable strategies for them. In this paper, Customer Lifetime Value (CLV) is used for this purpose. At first, CLV and its different calculating approaches are described. Then it used to analyze the customers of an online store called "Alakdolak". To calculate CLV, the probability of customer defection and the probability of buying three kinds of product named "low price", "medium price" and "high price" are calculated. Linear regression is used as a data mining technique for these models and verification and validation is done for each model. The validation process revealed that unlike the other models, "medium price model" does not reflect the actual behavior of customers truly.

Finally with calculating CLV, customers of the store are categorized into three different segments: strangers, barnacles, and true friends. The results showed that barnacles create most of the values of the company.

So company can increase the number of low price products to increase its profit. True friends spend more money in the store and have high loyalty. So the company should extend its interactions with these customers to increase the number of them, and decrease the number of strangers that buy just one time.

References

Bitran, G. R. and Mondschein S. V., (1996), Mailing Decisions in the Catalog Sales Industry. *Management Science*, 42(9), 1364-1381.

Bitterer, A., (2002), Customer relationship management: Application delivery strategies. *Meta Group white paper,* 15 March.

Blattberg, R.; Getz, G. and Thomas, J. S., (2001), *Customer equity: Building and managing relationships as valuable assets.* Boston, MA: Harvard Business School Press.

Blattberg, R. and Deighton, J., (1996), Managing marketing by the customer equity test. *Harvard Business Review*, 75(4), 136-44.

Colombo, R. and Jiang, W., (1999), A stochastic RFM model. *Journal of Interactive Marketing,* 13(3), 2-12.

Courtheoux, R., 1995, *Customer retention: How much to invest, research and the customer lifecycle*. 2nd edition, New York, NY: DMA.

Coussement, K. and Van den Poel, D., (2009), Improving customer attrition prediction by integrating emotions from client / company interaction emails and evaluating multiple classifiers. *Expert Systems with Applications*, 36(3), 6127-6134.

Cui, D. and Curry, D., (2005), Prediction in marketing using the support vector machine. *Marketing Science*, 24(4), 595-615.

Fader, P. S.; Hardie, B. G. S. and Berger, P. D., (2004), Customer-base analysis with discrete-time transaction data. Unpublished working paper.

Fader, P. S. and Lee, K. L., (2005), RFM and CLV: Using Iso-CLV curves for customer base analysis. *Journal of Marketing Research,* 42 (November), 415-30.

Friedman, J. H, (2003), *Recent advances in predictive (machine) learning*. Working paper, Department of Statistics, Stanford University, Stanford, CA.

Gupta, S.; Hanssens, D.; Hardie, B.; Kahn, W.; Kumar, V.; Lin, N.; Ravishanker, N. and Sriram, S., (2006), Modeling customer lifetime value. *Journal of Service Research*, 9 (2), 139-155.

Gupta, S. and Lehmann, D. R., (2003), Customers as assets. *Journal of Interactive Marketing*, 17(1), 9-24.

Gupta, S. and Lehmann, D. R., (2005), *Managing customers as investments*. Philadelphia: Wharton School Publishing.

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

Gupta, S. and Zeithaml, V., (2006), *Customer metrics and their impact on financial performance.* Working paper, Columbia University, New York.

Hansotia, B. and Wang, P., (1997), Analytical challenges in customer acquisition. *Journal of Direct Marketing*, 11(2), 7-19.

Hughes, A., (2005), *Strategic database marketing*. 3rd edition. New York: McGraw-Hill.

Hwang, H.; Jung, T. and Suh, E., (2003), An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry. *Expert Systems with Applications*, 26(2), 181-188.

Jackson, D. R., (1994), Strategic application of customer lifetime value in the direct marketing environment. *Journal of Targeting Measurement and Analysis for Marketing*, 3(1), 9-17.

Kecman, V., (2001), Learning and soft computing: Support vector machines. *Neural Networks and Fuzzy Logic Models*. Cambridge, MA: MIT Press.

Kumar, V., (2006a), CLV: *A path to higher profitability*. Working paper, University of Connecticut, Storrs.

Kumar, V., (2006b), *Customer lifetime value*. in Handbook of Marketing Research, Rajiv Grover and Marco Vriens, eds. Thousand Oaks, CA: Sage, 602-27.

Kumar, V.; Venkatesan, R. and Reinartz, W., (2006), Knowing what to sell, when to whom. *Harvard Business Review*, 84 (March), 131-37.

Menard, S., (1995), *Applied logistic regression analysis*. Series: Quantitative Application in the Social Science, Thousand Oaks, CA: Sage.

Pearson, S., (1996), *Building brands directly: Creating business value from customer relationships*. London: MacMillan Business.

Reichheld, F. F., (1996), *The loyalty effect*. Cambridge, MA: Harvard Business School Press.

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

Reinartz, W. and Kumar, V., (2003), The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing,* 67(1), 77-99.

Reinartz, W.; Thomas, J. and Kumar, V., (2005), Balancing acquisition and retention resources to maximize customer profitability. *Journal of Marketing*, 69(1), 63-79.

Roberts, M. L. and Berger P. D., 1989, *Direct marketing management*. Englewood Cliffs, NJ: Prentice-Hall.

Rosset, S.; Neumann, E.; Eick, U. and Vatnik, N., (2003), Customer lifetime value models for decision support. *Data Mining and Knowledge Discovery*, 7, 321-339.

Rust, R. T.; Lemon, K. N. and Zeithaml, V. A., (2004), Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing*, 68(1), 109-127.

Rust, R. T.; Ambler, T.; Carpenter, G.; Kumar, V. and Srivastava, R., (2004), Measuring marketing productivity: Current knowledge and future directions. *Journal of Marketing*, 8(4), 76-89.

Schmittlein, D. C.; Morrison, D. G. and Colombo, R., (1987), Counting your customers: Who they are and what will they do next? *Management Science*, 33(1), 1-24.

Sheth, J. N.; Parvatiyar, A. and Shaines, N., (2002), *Customer relationship management: Emerging concepts, tools and applications*. 2nd edition, New Dehli: Tata McGraw-Hill.

Sublaban, C. S. Y. and Aranha, F., (2008), Estimating cellphone providers' customer equity. *Journal of Business Research,* 62(9), 891-898.

Tarokh, M. J. and Sekhavat , A., A., (2006), LTV model in consultant sector, Case Study: Mental health clinic. *Behaviour & Information Technology*, 25(5), 399-405.

Thomas, J., (2001), A methodology for linking customer acquisition to customer retention, *Journal of Marketing Research,* 38(2), 262-68.

Thomas, J.; Blattberg, R. and Fox, E., (2004), Recapturing lost customers. *Journal of Marketing Research*, 41(1), 31-45.

Vapnik, V., (1998), *Statistical learning theory*. New York:Wiley.

Venkatesan, R. and Kumar, V., (2004), A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of Marketing*, 68(4), 106-25.

Villanueva, J.; Yoo, S. and Hanssens, D. M., (2006), *The impact of marketing-induced vs. word-of-mouth customer acquisition on customer equity*. Working paper, University of California, Los Angeles, Anderson School of Management.

Yoo, S. and Hanssens, D. M., (2005), *Modeling the sales and customer equity effects of the marketing mix*. Working paper, University of California, Los Angeles, Anderson School of Management.

Zeithaml, V.; Lemon, K. and Rust, R., (2004), Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing*, 68(1), 109-126.