

USING DISCRIMINANT ANALYSIS IN RELATIONSHIP MARKETING

Iacob Cătoi¹,
Mihai Țichindelean²,

ABSTRACT: The purpose of the present paper is to describe and apply discriminant analysis within a relationship marketing context. The paper is structured into two parts; the first part contains a literature review regarding the value chain concept and the dimensions it is built on, while the second part includes the results of applying discriminant analysis on several value chain dimensions. The authors have considered the client-company relationships of the gas-station market as proper for studying the differences between gas-stations based on relational dimensions, such as: perceived value and satisfaction.

Keywords: relationship marketing, discriminant analysis, value chain, perceived value, satisfaction

JEL Codes: M31, C31

Literature Review

Relationship Marketing – Brief History and Basic Concepts

The theory's purpose is to improve the scientific knowledge through systemized structures which are able to explain and predict phenomena (Hunt, Arnett, Madhavaram, 2006). Relationship marketing theory is considered an approach (Vargo, Lusch, 2004) or the definitive approach of understanding marketing phenomena (Harker, Egan, 2006). Due to its interdisciplinary character, relationship marketing is characterized by a conjunction of many theories which have influenced the marketing thought in several moments of time. Two scientific articles from business-to-business marketing have influenced the relationship marketing thought (Parvatiyar, Sheth, 1999). The first one (Adler, 1966) emphasized the importance of the marketing channel member for the companies' success. Thus, symbiotic marketing was defined as a resource or program alliance of two or more independent organizations with the purpose of increasing their market potential. The created symbiotic relationships between the channel members differed substantially from the classic transactional relationships between exchange partners.

The second article was written by John Arndt (1979) who asserted the idea of competitive markets erosion by considering in detail the inter-organizational system management and the political aspects of the companies' decision process. The author observed the companies trend of developing longtime exchange relationships with their main clients and suppliers and not discrete ones.

The ideas encompassed in the two mentioned articles were extended through American and European researches. In Europe, the relational approach of marketing was initiated in Sweden, in the sixties, by research efforts of the IMP Group (Industrial Marketing Purchase Group). This group was represented by researchers like Håkansson and Snehota (1995), Johansen and Mattsson (1985) and Kock (1991). Their studies were directed towards understanding industrial businesses (business-to-business marketing) from the network and internal interaction perspectives. Thus, a set of interaction between partners exists within a business network, which may consist of goods and information exchange, financial exchange or simple social actions. Due to this network complexity,

¹ Academy Of Economic Studies Bucharest, Romania, e-mail: iacob_catoi@yahoo.com

² Lucian Blaga University, Sibiu, Romania, e-mail: mihaitichi@yahoo.com

the role of marketing is unclear, therefore part-time marketers were defined as those employees which had not only their specific job responsibilities, but also marketing specific ones (employees of the research, logistic or customer relationships departments) (Gummesson, 1991). In the USA, researches which aimed the development of relationship marketing theory studied the factors which guided an exchange relationship between two partners (Levinthal, Fichman, 1988). Uzzi (1997) demonstrates that there are two categories of exchange relationships: (1) arm-length relationships characterized by a formal communication between the involved members and by the unique character of its happening and (2) embedded relationships defined through their social composition based on trust, information flow and common efforts in problem solving.

Relationship marketing – as an expression – was used for the first time in the year 1983 (Berry, 1983). Some authors consider that the best approach of understanding relationship marketing is by understanding service marketing (Berry, 1983); other authors propose an association of relationship marketing with network management (Håkansson, Snehota, 1995), partnership management (Hunt, Morgan, 1994) or integrated marketing communication (Lindberg-Repo, Grönroos, 2004). A complete definition for relationship marketing was developed by Grönroos (1990) who considered that relationship marketing is all about establishing, maintaining, enhancing and commercializing customer relationships (often but not necessary always long term relationships) so that the objectives of the parties involved are met. This is done by mutual exchange and fulfillment of promises. The relationship marketing theory includes a vast set of concepts which were developed and empirically tested through a variety of quantitative researches. Two main approaches of understanding relationship marketing theory have emerged out of this broad research: lifecycle of client relationships and value chain.

Client relationships or exchange relationships can be understood as a complex structure formed out of heterogeneous parts (variables) which through interaction result in a coherent behavior of the whole. One of the most important concepts of relationship marketing is the value chain which is based on variables which interact one with each other. By analyzing these variables and the effects of their interaction, relevant information regarding the relationship marketing decision process can be obtained (figure 1).

The value chain structure includes the following four elements:

- a. the company's activities (company's input)
- b. psychological effects of the clients (effects of company's activities)
- c. behavioral effects of the clients
- d. the company's economic result (company's output).

If considered individually, the value chain variables can be grouped into two categories: company specific variables (which can be controlled and modeled by the company) and client specific variables. For the company, the logic of the value chain is to develop those relationship marketing actions which generate psychological effects within the clients' minds, which further determine behavioral effects that are beneficial for the company. Every element of the value chain is influenced by moderators that can be endogenous or exogenous for the company. Thus, the value chain system can be considered a cybernetic system with relationship marketing measures – as entries – and economic results as outcome. The processes (psychological and behavioral effects) that occur between the entry-outcome cannot be clearly determined or measured, but only deduced.

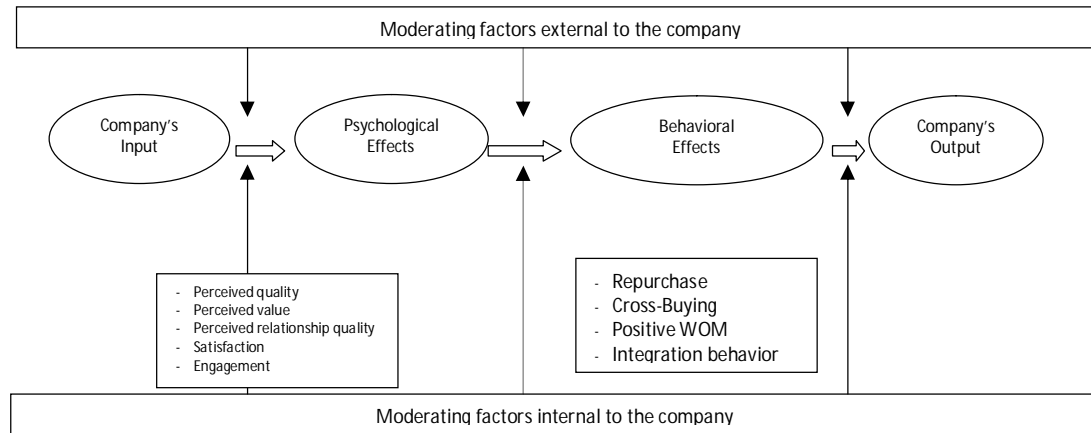


Figure no. 1 – Value Chain Concept (Bruhn, 2009)

Multiple linear discriminant analysis

The multiple linear discriminant method's objective is to discriminate (or differentiate) among the groups of one categorical variable based on a set of metric variables. This method is a statistical multiple analysis technique throughout which the linear relationship between a categorical dependent variable and one or more algebraic linear combinations between several independent variables measured on metric scales (Cătoi, 2009) is estimated. If the dependent, categorical variable is a dichotomous variable (is characterized by only two groups), the method is named linear discriminant for two groups; otherwise it is called multiple linear discriminant method.

By reviewing the marketing research literature, the following method objectives were identified (Malhotra, 2010):

1. identification of those discriminant functions based on algebraic linear combinations of the independent variables which best differentiate the independent variable's groups
2. examination if the differences between the dependent variable's groups are due to the individual predictors (independent variables)
3. determination of which predictor variables contribute to most of the intergroup differences
4. classification of cases to one of the groups based on the values of the predictor variables
5. evaluation of the accuracy of classification.

As every statistical multivariate analysis technique, the multiple linear discriminant method assumes the following hypotheses:

1. the used observation are based on random sampling
2. the independent variables are measured on metric scales, while the dependent variable is measured on nominal ones
3. every predictor (independent variable) is normally distributed
4. for every predictor, every observation is part of only one category of the dependent variable
5. the variance-covariance matrix should not differ significant between the dependent variable's groups.

The authors will use predictors instead of independent variables and factor instead of dependent variable to avoid future words agglomeration. Thus, the purpose of the multiple linear discriminant technique is to find that linear algebraic combination between the predictors that best discriminates the factors' groups.

The method's procedure is based on the matrix HE^{-1} which through the decomposition

$$HE^{-1}u_{n1} = \lambda_1 u_{n1} \tag{1}$$

results in identifying Eigenvectors and Eigenvalues. An Eigenvector represents that specific vector of a square matrix which multiplied (on the left) with a constant results in the same square matrix multiplied (on the right) with the Eigenvector. Every Eigenvector has a length, its Eigenvalue. Eigenvector identification is done by simple matrix calculus, as follows:

Square matrix $HE_{2 \times 2}^{-1} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$, Eigenvector $u_{21} = \begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}$ and Eigenvalue λ_1 are considered, then equation 1 results in:

$$HE_{2 \times 2}^{-1}u_{21} - \lambda_1 Iu_{21} = 0 \tag{2}$$

or

$$\begin{pmatrix} x_{11} - \lambda_1 & x_{12} \\ x_{21} & x_{22} - \lambda_1 \end{pmatrix} u_{21} = 0 \tag{3}$$

By equaling the matrix determinant with 0, Eigenvalue λ_1 can be determined. Based on its value, the coordinates of the Eigenvector $u_{21} = \begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}$ can be computed.

The elements of the matrix HE^{-1} contain the variation of the predictor set caused by their variation between the factor's groups related to their variation within the factor groups. Through the decomposition, the information of the initial matrix is not lost. The Eigenvector's coordinates represent the unstandardized canonical coefficients of the discriminant functions; they can be interpreted as partial correlations between every predictor and the identified discriminant functions.

For the considered theoretical example, one Eigenvector is formed with the following expression:

$$F_1 = u_{11}X_1 + u_{21}X_2 \tag{4}$$

Every Eigenvector has a length which is called Eigenvalue. The identified Eigenvalues represent the discriminating power of the identified discriminant functions. Thus, the bigger these Eigenvalues are, the more can the factor groups be differentiated by the discriminant functions. The statistical validity of the discriminant functions is verified through a statistical test called Wilks' Lambda test.

Research methodology

The research purpose is to identify if the client's perceived value and felt satisfaction discriminate the considered companies. To achieve this purpose, the authors used a survey as research method and a questionnaire as data collecting instrument. In accordance with the research purpose and method, the used information source has had the following characteristics:

Table no.1

Classification of the used information sources (Cătoi, 2009)

No.	Criteria	Source type	Explanations
1.	Source origin	External source	The used external source is represented by gas station customers who live in Sibiu county.
2.	Information type	Primary information	The data gathered from the respondents is primary data because it was gathered and analyzed for the first time within this research.
3.	Source identity	Individual	The individual (gas station customers) represents the bases for the present research.

4.	Cost of the gathered information	Paid source	For the present research the needed data has been collected by using both the offline and the online environment. For the online environment, a specialized survey site was used (http://freeonlinesurveys.com/) which has been a cost of 10 USD monthly.
-----------	----------------------------------	-------------	---

A conceptual and operation definition of the used variables is presented in the following table:

Table no.2

Conceptual and operational definition of variables

Independent variable – Perceived value								
Conceptual definition: Perceived value is the ratio between the benefits and sacrifices which are perceived by the customer after the consumption experience.								
Operational definition: For the measurement of the perceived value variable, the authors have used an interval scale (Likert type) with seven categories where 1 has represented total disagreement and 7 total agreement. Based on your last consumption experience with the company, circle the number which best describes your agreement level regarding the following:								
I had a good mood after my last consumption experience with the company.								
Totally disagree			Indifferent					Totally agree
1	2	3	4	5	6			7
Independent variable – Felt satisfaction								
Conceptual definition: Satisfaction is the pleasure or disappointment felt by a client after his consumption experience when he compares his initial expectations with the perceived performance of the company (Kotler, Keller, 2008).								
Operational definition: Satisfaction was measured using a performance scale (Rust, Zahorik, 1993), interval type with nine categories.								
Please circle the number which best represents your general satisfaction level felt within your last consumption experience.								
Very unsatisfied								Very satisfied
1	2	3	4	5	6	7	8	9
Dependent variable – Company								
Conceptual definition: The fuel company which was lastly patronized by the client.								
Operational definition: A nominal was used to measure the mentioned variable.								
Fill in the blank space with the name of the company you have lastly patronized								

The researched population was represented by the gas station customers who live within the Sibiu county area. Thus, three selection variables were used: possession or property of a car, using the car in personal purpose minimum four times a week, and a minimum age of 18 years. For this

research, the investigated population is the same with the survey unit, the individual which has offered the needed information according to his client experience, respectively.

Data was collected through a direct, simple cross-sectional research, ad-hoc survey type. A questionnaire was applied online and offline in the period September 2012 – March 2013. A probabilistic, simple random, sampling technique was used to select the respondents for the needed sample, thus, every individual of the defined population has had the same chance of being part of the sample. The final sample size was computed by using the formula below:

$$n = \frac{t^2 * p * q}{e^2} \tag{5}$$

where: - n – sample size

- t – coefficient value according to the desired significance level
- p – probability that the sample’s components have a certain attribute
- q - probability that the sample’s components do not have a certain attribute (q = 1 - p)
- e – error term.

The present research has used a 95% significance level and an error term of ± 6,3%. The authors have chosen the following probabilities: p = 0,5 and q = 0,5 because it was difficult to find a specific attribute which would have made a clear difference between the members and non-members of the investigated collectivity. Based on these data, the final sample size is:

$$n = \frac{1,96^2 * 0,5 * 0,5}{0,063^2} = 241 \tag{6}$$

and has the following structure:

Table no. 3

Sample structure based on socio-demographic and behavioral variables

Socio-demographic and behavioral variables	Variable categories	Percentage (%)	Absolute value
Age segment	Between 18 – 30 years	60,8	149
	Between 31 – 40 years	16,7	41
	Between 41 – 50 years	11,8	29
	Over 51 years	10,6	26
Gross income obtained in the last month	Under 1000 Lei	15,1	37
	Between 1000 – 2000 Lei	36,3	89
	Between 2001 – 3000 Lei	21,2	52
	Between 3001 – 4000 Lei	9,4	23
	Over 4000 Lei	17,9	44
Education level	High school	2,9	7
	Professional school	0,4	1
	Bachelor Degree	35,9	88
	Master degree	33,9	83
	Doctoral degree	21,6	53
	Post-doctoral degree	5,3	13
Car usage within a normal working day	None	3,7	9
	Once	13,5	33
	Twice	38,8	95
	Three times	18,0	44
	Four times	11,0	27
	More than four times	15,1	37
Car usage in weekends (Saturday and Sunday)	None	3,3	8
	Once	19,6	48
	Twice	34,3	84
	Three times	9,4	23
	Four times	10,6	26
	More than four times	22,9	56

IBM SPSS V.20 was used to codify and analyze the gathered data.

A first part of the analysis includes the computation of the predictors' variance and covariance for every category of the considered factor (table 3)

Table no.4

Predictor's variances and covariances within the factor categories

Factor (company)	Predictor	Satisfaction	Perceived value
Omv	Satisfaction	2,159	1,031
	Perceived value	1,031	1,524
Petrom	Satisfaction	2,648	1,183
	Perceived value	1,183	1,583
Mol	Satisfaction	2,082	1,057
	Perceived value	1,057	1,222
Rompetrol	Satisfaction	1,869	1,601
	Perceived value	1,601	2,458
Total	Satisfaction	2,327	1,219
	Perceived value	1,219	1,684

The covariances are positive within each of the four factor categories, meaning a positive relationship between perceived value and felt satisfaction (felt satisfaction varies in the same sense as perceived value). The variance-covariance matrix equality is tested by applying the Box test which verifies the null hypothesis that the variance-covariance matrix does not differ throughout the factor categories. A significance level of 0,474 determines the acceptance of the null hypothesis (table 4)

Table no. 5

Results of the Box test

Computed value of the Box test	8,845
F	0,957
Degrees of freedom	9
Degrees of freedom	35284
Significance level	0,474

The elements of the matrix HE^{-1} can be decomposed in specific Eigenvectors and Eigenvalues, as follows:

$$HE^{-1}u_{n1} = \lambda_1 u_{n1} \quad (7),$$

where:

HE^{-1} – resulted by multiplying the model matrix H with the inverse of the residual matrix (E^{-1});

u_{n1} – the first Eigenvector obtained through the decomposition characterized by n coordinates;

λ_1 – the first Eigenvalues specific for the first Eigenvector.

In our case, $n = 12$, thus, the Eigenvectors are represented in a bidimensional space. The presented Eigenvectors can be understood as vectors which best discriminate (in the bi-dimensional space formed by the two predictors – satisfaction and perceived value) the categories of the considered factor. The authors will use the terms discriminant vector instead of Eigenvector. For every Eigenvector there is an Eigenvalue which, mathematically, represents a resizing of the discriminant vector, and, statistically, the intensity in which the discriminant vector differentiates

the factor categories. The following discriminant vector coordinates are obtained through decomposition:

Table no. 6

Canonical coefficients of the discriminant vectors (or functions)

Predictors	Discriminant vectors	
	1	2
Satisfaction	-0,094	0,825
Perceived values	0,864	-0,514
(Constant)	-3,385	-3,208

Table no. 7

Computed Eigenvalues for the two discriminant functions

Discriminant function	Eigenvalue	% of variation	Cumulated variation	Canonical correlation
1	0.098	90,3	90,3	0,298
2	0.010	9,7	100,0	0,102

The statistical significance of the computed Eigenvalues is tested through the Wilks' Lambda statistic which is based on the χ^2 distribution (table).

Table no.8

Results of Wilks' Lamba test

Testing of the discriminant function	Wilks' Lambda	χ^2 computed value	Degress of freedom	Significance level
1 through 2	0,902	19,164	6	0,004
2	0,908	1,923	2	0,382

The computed values of the Wilks' Lambda test are interpreted as the variance of the predictors which is unexplained by the discriminant vectors. For our case, the first discriminant function explains only $1 - 0,902 = 0.098$ of the total predictors' variance. This result is confirmed by the values of the canonical coefficients of the discriminant functions. Thus, the canonical correlation between the first discriminant function and its predictors is 0,298 (table xxx); by squaring this value 0,088 is obtained which is similar with $1 - 0,902 = 0,098$. Based on the significance level, only the first discriminant function is further taken in consideration.

As mentioned, the discriminant functions can be written as algebraic linear combinations of the initial predictors:

$$F_1 = (-0,094)S + 0,864PV \tag{8}$$

$$F_2 = 0,825S + (-0,514)PV \tag{9}$$

The coordinates of the discriminant vectors are known as canonical coefficients of the discriminant functions. They can be interpreted as the relative contribution of the initial predictor for the identified discriminant functions (they can be interpreted as partial correlation coefficients between the discriminant function and the initial predictors). By analyzing the first discriminant function, it can be noticed that satisfaction has a reduced influence (-0,094) on the discriminant function, while the perceived values has a major one (0,864). Thus, the created latent dimension (first discriminant function) is based highly on the perceived value variable.

After the canonical coefficients have been multiplied with the standard deviations of the predictors, the values of the standardized canonical coefficients are obtained:

Table no.9

Standardized canonical coefficients of the discriminant functions

Predictors	Discriminant functions	
	1	2
Satisfaction	-0,142	1,248
Perceived value	1,080	-0,642

The following figure shows how the two discriminant functions differentiate the four categories of the considered factor.

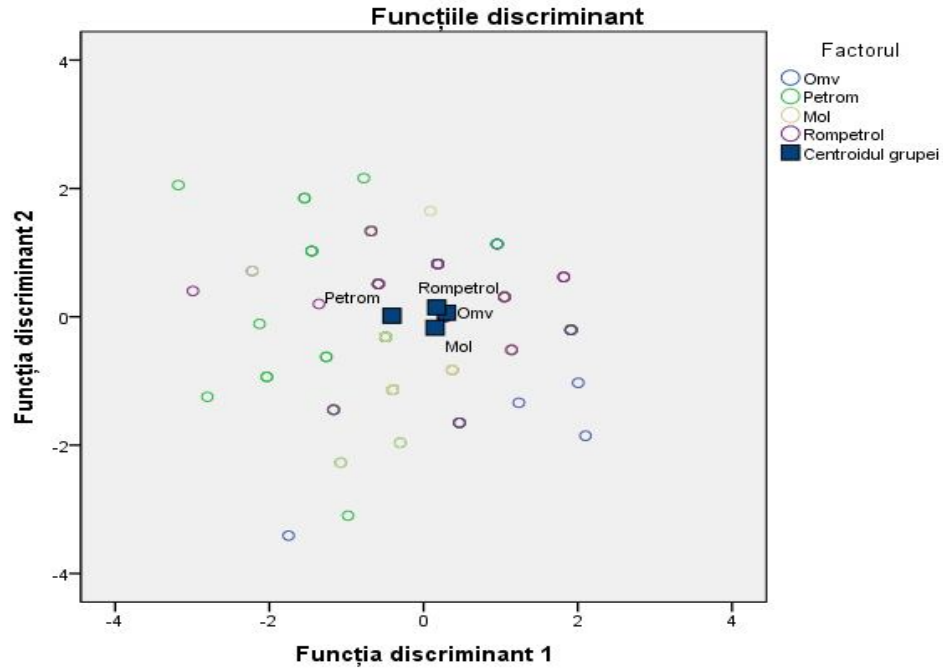


Figure no. 2 – Graphical representation of the groups’ centroids computed according to the identified discriminant functions

Only the first discriminant function is interpreted because of its statistical significance. Thus, this function groups the four factor categories in two groups: the first group contains the Petrom company, while the second group the other three companies: Rompetrol, Mol and Omv. The blue squares represent the centroids of the factor categories computed according to the two discriminant functions. Based on table 9, it can be concluded that the small discriminating effect on the factor categories is due mainly to the perceived value variable.

Conclusions

Based on the obtained results, the authors formulate the following four conclusions:

- (1) The initial predictors (satisfaction and perceived value) have formed two discriminant vectors, out of which only one was statistically significant
- (2) Based on the statistical significant discriminant vector, a discriminant function was built through an algebraic linear combination of the initial predictors. In a relationship marketing context, this new dimension can be considered a higher-order variable (for example: perceived relationship quality) which is determined by the client’s perceived value and satisfaction

- (3) The function has a reduced discriminating effect (0,098) on the companies, but a statistical significant one. Thus, Petrom company can be considered different from the other three companies (Omv, Rompetrol and Mol) based on the developed discriminant function
- (4) The upper mentioned differentiation is mainly due to the high influence of the perceived value on the discriminant function.

References

1. Adler L., 1966. *Symbiotic Marketing*, Harvard Business Review, Vol. 44
2. Arndt J., 1979. *Toward a Concept of Domesticated Markets*, Journal of Marketing, Vol. 43
3. Berry L.L., 1983. *Relationship Marketing, Emerging Perspectives on Service Marketing*, American Marketing Association.
4. Bruhn M., 2009. *Relationship Marketing – Das Management von Kundenbeziehungen*, 2te Auflage, Vahlen Verlag, München
5. Cătoiș I., (coord), 2009. *Tratat de cercetări de marketing*, Ed. Uranus
6. Håkansson H., Snehota I., 1995. *Developing Relationships in Business Networks*, Routledge, London
7. Field A., 2005. *Discovering Statistics Using SPSS*. Second Edition, Sage Publishing
8. Grönroos Ch., 1990. *The Marketing Strategy Continuum: Towards a Marketing Concept for the 1990s*, Management Decision, Vol. 29, No. 1
9. Gummesson E., 1991. *Marketing Revisited: The Crucial Role of the Part-Time Marketer*, European Journal of Marketing, Vol. 25, No. 2
10. Harker M.J., Egan J., 2006. *The Past, Present and Future of Relationship Marketing*, Journal of Marketing Management, Vol. 22.
11. Hunt S.H., Arnett D.B., Madhavaram S., 2006. *The Explanatory Foundations of Relationship Marketing Theory*, Journal of Business & Industrial Marketing, Vol. 21, No. 2
12. Johansson J., Mattsson L.-G., 1985. *Marketing Investments and Market Investments in Industrial Networks*, International Journal of Research in Marketing, Vol. 2
13. Kock S., 1991. *A Strategic Process for Gaining External Resources through Long-Lasting Relationships: Examples from Two Finnish and two Swedish Industrial Firms*, Helsingfors: Svenska Handelshogskolan
14. Kotler Ph., Armstrong G., 2010. *Principles of Marketing – Thirteenth Edition*, Pearson Prentice Hall
15. Levinthal D.A., Fichman M., 1988. *Dynamics of Interorganizational Attachments: Auditor-Client Relationships*, Administrative Science Quarterly, Vol. 33
16. Lindberg-Repo K., Grönroos C., 2004. *Conceptualising Communications Strategy from a Relational Perspective*, Industrial Marketing Management, Vol. 33
17. Malhotra N.K., 2010. *Marketing Research. An Applied Orientation. Sixth Edition*, Pearson Education Inc
18. Vargo S.L., Lusch R.F., *Evolving to a New Dominant Logic for Marketing, 2004*, Journal of Marketing, Vol. 68
19. Parvatiyar A., Sheth J.N., 1999. *The Domain and Conceptual Foundations of Relationship Marketing in Handbook for Relationship Marketing*, Sage Publications, Thousand Oaks
20. Uzzi B., 1997. *Social Structures and Competition in Interfirm Networks: The Paradox of Embeddedness*, Administrative Science Quarterly, Vol. 42, Issue 1.