# THE OPTIMAL LOCATION OF A RETAIL STORE

by

Shyrkova Kateryna

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Arts in Economics

National University "Kyiv-Mohyla Academy"
Economics Education and Research Consortium
Master's Program in Economics

2005

Approved by – _____
Ms.Svitlana Budagovska (Head of the State Examination Committee)

_____

_____

_____

Program Authorized
to Offer Degree _____Master's Program in Economics, NaUKMA_____

Date _____

National University "Kyiv-Mohyla Academy"

Abstract

# THE OPTIMAL LOCATION OF A RETAIL STORE

by Shyrkova Kateryna

Head of the State Examination Committee: Ms.Svitlana Budagovska,
Economist, World Bank of Ukraine

The most important issue in the investment decision for the retail market is location, which draws a considerable attention to the problem of optimal location both for the scientists and businessman. One of the models here is the maximum capture model, which is a model that offers a technique of the location choice based solely on the distance. This paper presents the application to the Ukrainian market of the new methodology for deciding which store attributes should be included into the maximum capture model, when it is applied to the retail market. The methodology represents designing of a survey with its further examination by the factor analysis and regression methods.

TABLE OF CONTENTS

# LIST OF FIGURES AND TABLES

# ACKNOWLEDGMENTS

The author wishes to express gratitude to her adviser, Dr. Volodymyr Bilotkach, for his support and valuable comments during the whole period of work. The special words of thankfulness are devoted to the Research Worksop professors, Dr. Tom Coupé and Polina Vlasenko, for their useful ideas and helpful comments from the very first research proposal.

# GLOSSARY

**MaxCap** Maximum Capture Problem.

**MCI** Multiplicative Competitive Iteration model

**GLM** Generalized Linear Model

**OLS** Ordinary Least Squares

**ML** Maximum Likelihood

**HAC Covariance** Heteroskedasticity and Autocorrelation Consistent Covariance

**DW-statistics** The Durbin-Watson statistics

*Chapter 1*

# INTRODUCTION

These days the retail industry in Ukraine is considered to be one of the fast growing. But this industry is not explored. The reasons for that are the following.

- Modern format stores are relatively 'young' in Ukraine.
- The development of a local retail chain began not so long ago; therefore, the existing local stores are not experienced in optimal location choice.
- The well-established retail chains are not familiar with the particularities of the Ukrainian retail markets and need to take into account the local conditions.
- Retail market in Ukraine exhibits rapid development.

As an evidence for it, we can consider the Global Retail Development Index (GRDI), which is supported by A.T. Kearney. This index ranks emerging countries on the basis of economic and political risk, the level of retail situation, and the difference between GDP growth and retail growth. In 2004 Ukraine obtained 11-th place among 30 developing countries (20-th in 2003) and directly approached the group of countries where it is advisable to go immediately.

Therefore, one of the hottest topics in theory and practice of a retail chain is the location of the stores. In a market economy, where one of the main driving forces is rivalry, the choice of a location for retail and service firms within the broader context of a well-formulated competitive strategy is crucially important.

To deal with competitor policies, retail managers can change such factors as price, variety and quality of the product and service; however, the same cannot be said

about the location, which represents a fixed one-time investment of a unique, unchangeable nature (Colomé and Serra, 2003). Because the issue of convenience is of great importance to today's consumers, a retail store can prosper or fail solely based on its location.

The issue of the optimal location of the retail stores have been vigorously researched during the past century. At present there exist several key areas in the literature on the store location, competitive location literature and store-choice literature. The former concentrates on the problem of optimal location of the firms that compete for clients in space. The latest examine the key variables that influence a consumer when he makes a decision where to shop as well as interaction between the variables.

One of the recent trends in area is encompassing of the store-choice attributes in the competitive location models. One of the successful examples is the work of Colomé and Serra (2003). Their paper describe a new approach for determination which store characteristics should be included in the new Maximum Capture Model and how they ought to be reflected using the Multiplicative Competitive Interaction model.

The goal of this study is the application of the above-mentioned methodology for the Ukrainian retail market when deciding which store attributes should be included into the maximum capture model. The study is conducted through the several steps. The first step involved design of a survey for collecting of the information about the consumer preferences. The second step is examination of the survey data using the factor analysis technique. This procedure implies determining of the surrogate variables, where each variable represents the corresponding group of attributes of consumer choice. On the third step the surrogate variables obtained on the previous stage are incorporated in to the MCI model. The final step is calibration of the MCI model using the regression

techniques. On this stage the significant variables for the study and the sensitivity parameters are determined. The result of the study is incorporation of the calibrated MCI model into the maximum capture problem.

The thesis outline is as follows. Chapter 2 presents a review of the literature on the area of research. Chapter 3 described the methodology and the description of the data. Chapter 4 focuses on the model estimation and the empirical results. Chapter 4 provides the conclusions with the discussion of the limitation of the analysis and further development.

*Chapter 2*

# LITERATURE REVIEW

There is a great number of papers on the retail industry worldwide. One of the main problems on which the researchers focus their attention, is the problem of the location of the stores. Since the location problem is particularly important for both the theorists of the economics and the businessmen, there is substantial theoretical and empirical literature on the subject.

The issue of practical application of the store location models has become extremely popular with the development of retail giants, such as Tesco, Wal-Mart and others. However, the theoretical concept of the location model has been developing from the beginning of the XX century and gains the particular attention in the second half of the XX century.

The theoretical foundations of the issue of the firm's optimal location within the given area were established in the famous and widely cited paper "Stability in Competition" by Hotelling (1929). His research focuses on the problem of location in the competitive environment. In the paper the author examines two competing firms in a linear market with uniformly distributed consumers.

Nevertheless, in the mid-XX century the majority of retail chains were applying not the theoretical models of different approaches to the spatial analysis framework, but rather simple techniques of location choice (Clarke, 1998). The most popular among those techniques are 'gut feeling', 'checklist' and 'analogue'. The surveys of retail chain's location planning methods show that 'gut feeling'

and ‘checklist’ approaches were among the most popular up to the late 1980s (Simkin ,1990)

Gut feeling is the method that relies on the decision of a senior manager, who visits different locations and obtains some ‘gut feeling’ about the visited places. Although some researches, such as Davies (1977), claim the expert decisions to be reliable, ‘gut feeling’ technique obviously posses serious drawbacks due to high degree of subjectivity and time-consuming issues.

Among other simple methods there is a checklist approach, which involves procedures to measure the potential of existing centres and analysis of catchments area in terms of population structure (Clarke, 1998). The method includes incorporating as much information about the given neighbourhood in the model as possible. Based on the analysis the retailer obtains a rank of the possible locations.

The basic approach to the analogue techniques involves forecasting of the potential sales of a new store based on the results of similar stores. The analogy is drawn on the basis of physical, locational and trade area characteristics. The comparison can be done through the regression techniques. One of the extensions of the analogue techniques is the so-called "follow my leader" or parasitic approach. That is to put the store in the location, were the anchor retailers already have put their stores.

The multiple regression model works by defining the dependent variable such as store turnover, which can be considered as aggregate measure of the effect of consumer shopping behaviour, which is store performance (Clarke, 1998). Then it correlates the dependent variable with the set of explanatory variables. Coefficients are calculated to weigh the importance of each independent variable in explaining the variation in the set of dependent variables. The model can be written as:

$$Y_i = a + b_1 X_{1i} + b_2 X_{2i} + b_3 X_{3i} + ... + b_m X_{mi} + ...$$

where Yi is turnover (the dependent variable) of store i, $X_{mi}$ are independent variables, $b_m$ are regression coefficients and *a* is the intercept term.

There are several variables that need to be considered during the research. The size and importance of the centre, which can be measured comparatively easily by common floorspace statistics obtainable from local authorities or private sector organizations. Such aggregate statistics can be broken down by type of retailer present in order to determine the 'quality' of that centre (Clarke, 1998). Relevant store outlet characteristics are also the store or chain image, its retail format and strategy.

A number of papers focus their attention on the determining the appropriate retail environment to achieve required level of store performance (Kumar and Karande, 2000). The research is concentrated on the internal store environment such as number of checkout counters per square foot of area, number of non-grocery items sold, whether a store has banking facilities, and whether a store is open 24 hours. Among the external factors are total number of households in the area and the geographic location.

The importance of the store format is examined by Campo and Gijsbrechts (2004). The issue that they address is wether retailer should adopt micro-marketing strategy in different store formats and what are the ways of adopting such a strategy. The researchers investigate the conditions that are favourable for such micro-marketing strategy depending on the store format. Another focus of the studies is how the model of adjustment depends on the store format.

Another important problem within the retail location choice is whether to build a supermarket within a city or outside the city area. One of the researchers in the area, Risto Murto from The Nordic Research Network on Modelling Transport

Land (2001) focuses on the transportation problem within the optimal location choice. The author presents a model for the transport, which can measure changes in traffic network from different land use alternatives. This model includes a logit and forecast model for shopping trips and a model for goods transport.

The location research is also considered a useful tool in examining other economic phenomena, such as mergers of several retail chains. An interesting example of such work is Birkin (2000), who focuses on possible application of the spatial analysis to the merger problem and introduces a framework for improving the process of merge in retail area.

Recent reviews of the store location techniques used in the latest time suggest that there is a considerable shift towards more sophisticated models (Clarkson et. al., 1996). There are several major directions in the literature on the store location, which are depicted in the recent surveys: competitive location literature and store-choice literature (Yrigoyen and Otero, 1998; Colome and Serra, 2003).

Competitive location literature.

This literature addresses the issue of optimally locating firms that compete for clients in space. The key model here is Maximum Capture Problem (MaxCap) developed by ReVelle (1986), which is one of the most popular facility location models, that has important theoretical findings and practical implications. The objective of MaxCap is to establish a set of finite number of facilities in order to maximize the total weight of customers, covered by those facilities. The problem was initially researched by Church and ReVelle (1974). One of the principal assumptions of the MaxCap is that a customer location is either fully covered or not covered at all. However, this assumption sounds unrealistic in the area of the retail facilities research.

A number of research papers consider the possibility of the partial coverage in the maximal cover location problem. Karasakal (2004) develops a notion of partial coverage, which is defined as a function of the distance of the demand point to the facility. In the model the demand point can be fully covered within the minimum critical distance, partially covered up to a maximum critical distance, and not covered outside of the maximum critical distance. Based on the revised model the author makes a computation and compare the results with the results of the general model, which demonstrate that including the partial coverage has significant affects on the solution of the problem.

Another aspect of partial coverage is presented in the paper 'The generalized maximal covering location problem' (Karasakal, et. al., 2002). Here the authors develop a generalized maximal cover location problem (GMCLP) by introducing a multiple set of coverage levels for each demand point, where the coverage level is a decreasing function of the distance to the closest facility.

Store-Choice literature.

Store choice models consider the process of store choice by consumer by examining the key variables that affect the consumer's preferences over the stores. The literature of the store-choice problems can be classified into three groups (Craig, et. al.,1984).

The first group involves descriptive-deterministic approach to the location, which represent techniques that rely on some normative assumptions. The most common assumption is regarding consumer spatial behaviour as patronizing the nearest outlet. This group involves empirical observation techniques based on observation of the market areas. It uses customer surveys to find out the geographic of the certain trade area. One of the most-cited papers in this group is 'Law of retail gravitation' (Reilly, 1929), where the author states that "the

probability that a consumer patronize a shop is proportional to its attractiveness and inversely proportional to a power of distance to it". Reilly spanned the new direction of research in spatial choice models called now 'Gravity models'. The best representatives of Gravity models research are Yrigoyen and Otero (1998): Reilly's model (1929), Converse's model (1949), Baumol and Ide's model (1956), Batty's model (1978) and Albadalejo's model (1995).

The second group represents revealed preference approach to the modelling. The models use past behaviour of the consumers to understand the dynamics of retail competition and the determinants of the consumer choice. The group includes spatial interaction model presented by Huff (1964), where he incorporates the probability of the consumer to shop in one of the possible stores. This probability is defined as decreasing with distance to the store and increasing with the size. Although the Huff's model is crucial among the retail-choice models, it involves only two explanatory variables, which obviously does not represent the full picture. The extensions of the Huff's model are Multiplicative Competitive Interaction Model (Nakanishi and Cooper, 1974) and Gautschi's model (Gautschi, 1981). Those models add to Huff's model additional variables that represent the attractiveness of the store.

The third group includes models based on direct utility assessment. The models estimate consumer utility functions from simulated choice data using information integration, conjoint or logit techniques (Louviere and Woodworth, 1983). The best model in the group is Ghost and Graig model (1983), which is based on game theory.

Nowadays, retail location area experiences a trend of encompassing the store-choice attributes in the competitive location models. One of the successful examples is the work of Colomé and Serra (2003). In the paper authors present a new approach for determination which store attributes should be included in the new Market Capture Model and how they ought to be reflected using the Multiplicative Competitive Interaction model, as the latest stands to reason that any retail location model should take into account the processes underlying consumers' choice of store. From a consumer perspective, choice of a store depends upon whether the store is price-orientated or not, as this point leads to expectations about relative price levels offered by different stores and other characteristics. The methodology the authors use involves the creation of a survey and its further analysis with the application of factor analysis and ordinary least squares techniques.

*Chapter 3*

# METHODOLOGY AND DATA DESCRIPTION

## 3.1 Methodology

The aim of this research is to develop the Maximum Capture Problem (ReVelle, 1986) by incorporating different attributes of the consumer choice into the model. The initial MaxCap model searches for the location of a fixed number of the retail outlets for the firm that is going to enter a spatial market. The model assumes that several competing firms already exist in the market. The spatial market in the model is represented by a network. Each node of a network corresponds to a segment of local market with a predetermined demand. As a result, the nodes limit the choice of location. The store is considered captured if it is the nearest store within the given area. The entering firm is maximizing the market capture by the location choice. The formal model set up is presented below.

$$MAXZ = \sum_{i \in I} \sum_{j \in J} a_i \rho_{ij} x_{ij} \quad \text{, subject to} \tag{3.1}$$

$$\sum_{j \in J} x_{ij} = 1, \forall i \in I$$

$$x_{ij} \leq x_{jj}, \forall i \in I, \forall j \in J$$

$$\sum_{j \in J} x_{jj} = 1$$

$$x_{ij} = \{0,1\}, x_{jj} = \{0,1\}, \forall i \in I, \forall j \in J$$

The parameters and variables here are:

*i,I* – Index and set of consumer zones

*j,J* – Index and set of potential locations for stores.

$J^B(\in J)$ – The set of actual location of the existing stores.

$d_{ij}$ – The network distance between consumers' zone *i* and a store in zone *j*.

$d_{ibi}$ – The network distance from node *i* to the closest competitor shop $b_i$ .

$a_i$ – demand of consumer zone *i*.

$$x_{ij} = \begin{cases} 1, \text{if consumers zone i assigned to node j} \\ 0, \text{otherwise} \end{cases}$$

$$x_{ij} = \begin{cases} 1, \text{if a shop of firm A operates at node j} \\ 0, \text{otherwise} \end{cases}$$

$$d_{ij} = \begin{cases} 1, \text{ if the } d_{ij} \le d_{ibi} \\ 0, \text{otherwise} \end{cases} \qquad (3.2)$$

In the framework of initial model, the probability of consumer from zone i to shop in zone j is determined solely by location. In my research, I will apply the methodology, which is derived from Colomé and Serra (2003), to incorporate the different attributes of consumer choice other than location into the model. The parameters $p_{ij}$ , which characterise the attributes of consumer choice will be determined using Multiplicative Competitive Iteration model (MCI). During the research I will follow several steps.

First step. At the beginning of the research I need to design a survey to investigate the behaviour of a consumer faced with a choice of store. The purpose of the survey is to provide information for the calibration of the model to the local situation. The calibration is done using the revealed preference approach, which uses past consumer behavior to model the future behavior of the consumers.

The target population for the research was supermarket consumers in Ukraine. The survey was conducted in Kiev on 17.02.2005 (Thursday) and 19.02.2005 (Saturday). The first part was introduced to capture the weekday consumer behavior, whereas the second date was to capture the weekend consumer preference.

The location at which the study was conducted is the Troeschina region of Kiev, namely, two supermarkets: Billa and Silpo. These supermarkets are the representatives of the well-established retail chains in the retail market of Ukraine. Billa is the foreign retail chain that came to the Ukrainian market, whereas the Silpo is local chain that was originally established in Ukraine. The total number of 154 consumers was questioned, from which 83 were interviewed in Billa and 71 in Silpo.

In the research I employ simple random sampling procedure during which the shoppers in the supermarket were interviewed. Additionally the random eerc students were questioned as the representatives of a specific group of the stores' consumers.

The questionnaire includes several parts. The survey that I used in my research is based on the survey of Colomé and Serra (2003), however it was modified to capture the local characteristics of the retail market. The first part of the questionnaire was designed to determine the relative importance of different supermarket attributes for the consumers. The consumers were proposed to rank the following attributes from 1 to 5:

- location
- low prices
- quality and range of the products
- opening hours
- speed of service

- quality of service
- supermarket design.

Here 5 is the score of the most important attribute and 1 is the score of the least important attribute.

The attributes were taken from Colomé and Serra (2003) and adopted to the local retail market situation.

The second part of the questionnaire is the most important as it was designed to reveal the valuation that consumers give to the attributes of the particular store. The valuation ranged from 1 to 10, where 1 is the lowest valuation and 10 is the highest valuation. The key attributes were congregated in groups. Each of the groups represented the broadly defined store attributes such as convenience, customer service, store design, product characteristics and prices.

Based on the results of the survey, I form the following variables for attributes of the store choice. The exact questions for the survey can be found in Appendix 1.

1) Convenience of the transport services. This block represents consumer valuation of the stores attributes, which are connected with the transport accessibility.

| | |
|---|---|
| tr_publ | Easy access by Public transport |
| tr_parc | Convenient parking |

2) Checkout and shopping assistance. This group describes the scores that were given by the consumers to the different customer service attributes of the particular stores.

| | |
|---|---|
| serv_speed | Fast checkout |
| serv_ass | Shopping assistants are courteous |
| serv_qual | Qualified assistance |

3) Supermarket design and physical facilities. This block was designed to reveal the attitude of the shoppers to the physical characteristics of the supermarket.

| | |
|---|---|
| des_crowd | Not crowded store |
| des_move | Easy to move around the store |
| des_find | Easy to find products |

4) Quality and Range of products. This cluster shows the valuation of the product characteristics by the consumers.

| | |
|---|---|
| prod_basis | Store has basic products |
| prod_varie | Good variety of products |
| prod_fresh | Fresh products |
| prod_qual | High quality products |

5) Pricing Policy. The block reveals attitude of the shoppers towards the price policy of the particular stores.

| | |
|---|---|
| price_low | Low price store |
| price_prom | Promotional offers |
| price_adv | Advertising of products |

6) Opening hours. The group shows valuation of the opening hours of the supermarkets by the consumers.

| | |
|---|---|
| op_late | Open late at night |
| op_early | Open early in the morning |

7) Location. The variable represents how the consumer evaluates the location of a particular supermarket. This variable characterizes s the whole cluster.

| | |
|---|---|
| loc | Well located |

Second step. At the next stage of the analysis I need to evaluate the obtained results, that is, to determine the mechanism of consumer valuation of the

attributes. In my research I follow Colomé and Serra (2003) method based on of Hierarchical Information Integration approach (Louvierre, 1984).

The idea of the approach is to make use of the assumption that, when consumers decide on the supermarket to shop in, they organize individual decisions attributes into groups or clusters. They evaluate each attribute in the cluster with further estimation of some aggregate property for each cluster to obtain the overall judgment. This approach employs factor analysis for categorizing the attributes estimated in the survey into specific factors.

Factor analysis is employed to examine the patterns of relationship among many dependent variables, aiming to discovering something about the nature of the independent variables that affect them, even though those independent variables may not be measured directly. The factor analysis will be used to explain the variance in the observed variables by means of fundamental latent factors. The model specification is the following: it assumes that *p* observed variables (denoted X) measured for each of the n subjects have been standardized. This can be written in the matrix notation:

$$X_{p\times 1} = A_{p\times m} \times F_{m\times 1} + \varepsilon_{p\times 1}$$

The $F_i$ are the elements of vector of m common factors, the $\varepsilon_i$ are the *p×1* vector of specific errors, and the $a_{ij}$ are the elements of factor loadings matrix. The $F_i$ have mean that is equal to zero and standard deviation that is equal to unity, the factors are normally assumed to be independent.

In my research I will perform factor analysis using Principal Factor Analysis (PFA). The choice of the method is determined by the fact that the PFA method always gives a unique solution (with the exception for some very special cases). It takes the original estimation of the common factor for each variable and replaces with them the diagonals in the correlation matrix. Then, the method constructs

the principal components and takes the first m loadings. In my analysis I will use varimax rotation, which maximizes the sum of the squared factor loadings across the columns. As a result, the variables will be enforced to load as highly as possible to the new factors. Applying this method I will be able to identify which factors are the most significant into each cluster of the survey.

Third step. The further analysis involves specification of the Multiplicative Competitive Iteration model. At this stage of my research I will substitute variables of the MCI model with factors that were found in the previous stage of the analysis.

The formal MCI specification is taken from the original version of Nakanishi and Cooper (1974), that is:

$$p_{ij} = \frac{\prod_{k=1}^{s} A_{kij}^{\beta_k}}{\sum_{j=1}^{m} \prod_{k=1}^{s} A_{kij}^{\beta_k}}, \quad \text{where} \tag{3.2}$$

$p_{ij}$ – The probability that consumers located at $i$ will make purchases in store $j$.

$A_{kij}$ – The $k$-th attribute that describes store $j$ attracting consumers from zone $i$.

$i$ – consumer zone index, where $i = \overline{1,n}$

$j$ – store index, where $j = \overline{1,m}$

$\beta_k$ – parameters that reveal sensitivity of the consumers to the store attribute and as a consequence influence the probability of purchasing in the particular store.

Forth step. The next step of the research is calibration of the model using the estimated parameters obtained through the survey. The calibration involves two parts. First, we need to determine which variables are significant for our study. We also need to estimate parameters $\beta_k$, which reflect the sensitivity of

consumers to the chosen store choice attributes. The significance of parameters will be obtained using the significance of the estimated coefficients in the resulting regression. The parameters $\beta_k$ will be directly estimated from the regression.

Following Nakanishi and Cooper (1974), we can calibrate our MCI model employing the regression method. The ordinary least squares regression on the log-transformed centred form of the equation has proven to give efficient and unbiased estimators if the disturbances are uncorrelated and homoskedastic. However, this was not the case with the dataset under consideration.

In the presence of autocorrelation and heteroskedasticity the coefficients still are unbiased and consistent. However, the estimates are no more efficient, that is the standard errors are incorrect and should not be used for statistical inference. In my research I will employ two procedures to deal with non-spherical disturbances problem.

The first procedure Heteroskedasticity and Autocorrelation Consistent Covariance (HAC covariance), which is the technique, proposed by Newey and West. That is a more general covariance estimator that is consistent in the presence of both heteroskedasticity and autocorrelation of unknown form.

The second method is Generalized Linear Model approach (GLM). In my study I will use the Maximum Likelihood (ML) regression technique with the unbiased sandwich variance estimator.

In the line of my analysis, I will transform the original equation (2) to the log-transformed centred form (3), following Colomé and Serra.

$$p_{ij} = \frac{\left(\prod_{k=1}^{s} A_{kij}^{\beta_k}\right)\zeta_{ij}^*}{\sum_{j=1}^{m}\left(\prod_{k=1}^{s} A_{kij}^{\beta_k}\right)\zeta_{ij}^*} \tag{3.3}$$

$$\ln\left(\frac{p_{ij}}{\hat{p}_i}\right) = \beta_k \cdot \sum_{k=1}^{s} \ln\left(\frac{A_{kij}}{\hat{A}_{ki}}\right) + \ln\left(\frac{\zeta^*_{ij}}{\hat{\zeta}^*_i}\right), \text{ where} \tag{3.4}$$

$\hat{p}_i = \left(\prod_{j=1}^{m} p_{ij}\right)^{1/m}$ - Geometric mean of the probabilities of consumers at zone *i* shopping at *m*-stores.

$\hat{A}_{ki} = \left(\prod_{j=1}^{m} A_{kij}\right)^{1/m}$ - Geometric mean of *k*-th attributes of m stores evaluated by consumers at zone *i*.

$\hat{\zeta}_i = \left(\prod_{j=1}^{m} \zeta_{ij}\right)^{1/m}$ - Geometric mean of the specification error term of *m* retail facilities.

One of the most important factors for the consumer in his choice of store is the location of the nearest store. It is generally approved fact that the consumer will be more inclined to shop in the nearest supermarket other things equal. For that reason, I will include in the model a dummy variable *near*.

$$\text{near}_{ij} = \begin{cases} 1, \text{if store j is the nearest store for a consumer} \\ 0, \text{otherwise} \end{cases}$$

The data on store closeness was also obtained from the survey of the consumer.

Although the information about the consumer attributes obtained through the survey is rather substantial, there exist other subjective factors of the consumer

choice that may not be covered by this research. Consequently, I will include into the model specification the dummy variable for the store, which will incorporate all the information about the store characteristics, which were not captured by the survey data.

$$near_{ij} = \begin{cases} 1, \text{if store is Billa} \\ 0, \text{if store is Silpo} \end{cases}$$

In order to obtain the estimation results, the model needs to have the specification ensuring that al least one consumer will shop in every given store. In the data set under consideration this condition is not violated.

*Chapter 4*

# DATA DESCRIPTION

Within the first part of the survey the consumers were asked to rank the main attributes of the supermarket. There were 7 dimensions to be ranked (from very important 5 to not important at all 1). The summary statistic consists of mean and standard deviations for the attributes and are presented in Table 1. The percentage of the highest score for each of the attributes is presented on Figure 1.

Table.1 Ranking of the Supermarket attributes.

| Attribute | Mean | Standard deviation |
|---|---|---|
| Location | 3.62 | 1.19 |
| Low prices | 4.22 | 0.99 |
| Quality and range of products | 4.2 | 0.77 |
| Hours of opening | 3.9 | 1.01 |
| Speed of service | 3.66 | 1.15 |
| Quality of service | 3.64 | 1.09 |
| Design | 2.45 | 1.13 |

The results of the survey point to low price image of a store as the most important attribute of the supermarket choice. As can be seen from the Table 1, more than half of the respondents give the highest importance score to the low price image. This result is very intuitive. As the purchasing capacity of the population in Ukraine is very low, the price remains the most important factor for the consumers. Furthermore, the Troeschina region where the survey was conducted is considered as on the regions with poorest population in Kiev.

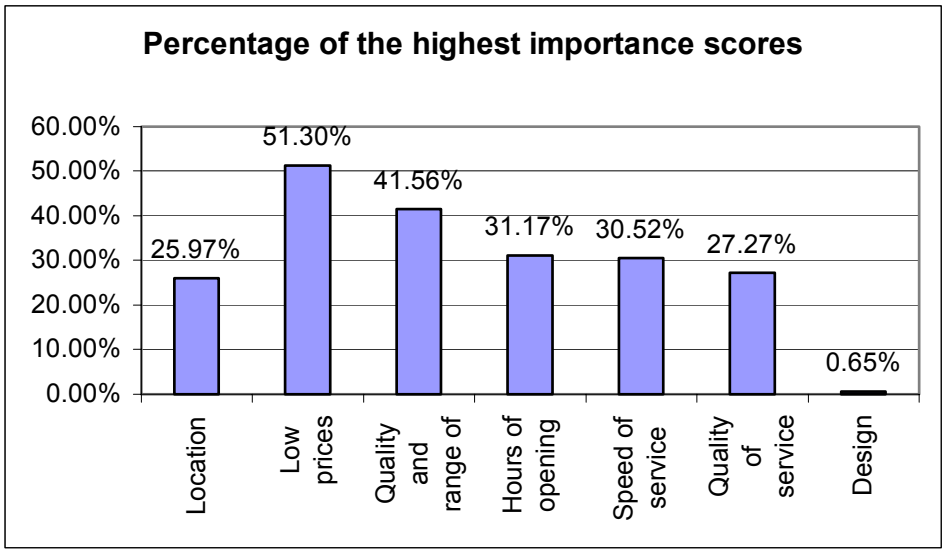


Figure 1. The percentage of the highest importance scores.

The next important factor is the quality and the range of product. To this attribute was given the highest valuation by more than 40% of the respondents. This result is also very intuitive and confirms the idea of high relevance of the product range and quality for the consumers. Among the other highly important attributes, we can single out hours of opening and speed of service. Surprisingly, location was named as attribute with highest valuation by only a quarter of the respondents. This result is difficult to explain. It may be obtained due to the data limitation. The least important factor is the design of the supermarket, which is also very intuitive result.

We can compare the result obtained for Ukraine with the results for Spain and Great Britain consumers. Those results are taken from the research paper of Colomé and Serra (2003). The authors found that the most important attribute for the Spanish shoppers is the convenience. The next factor by importance for the Spanish survey is the quality and range of product. However, the British consumers have evaluated the quality of products as the attribute with the highest

importance. The price image of a supermarket has received lower ranking for both countries.

As we can see, the quality and range of product is one of the most important factors for the consumers in all the countries, which is very intuitive result. However, the shoppers in Ukraine are different from the western consumers in the fact that they put the low price image of a store to the highest place. This fact confirms the idea of different shopping behaviour of Ukrainian consumers due to the low living standard.

Another piece of relevant information can be also drawn from the standard deviations of the attributes ranking. We can observe low deviation for the items like hours of opening, quality and range of products and the low price policy. This result could be expected, since the deviation in the attributes depends to some extent on the value of the mean. For the highly valuable attributes we expect to have high factor scores in the majority of questionnaires, which imply low deviation from the mean. For the less valuable attributes the evaluation is more individual and therefore it is subject to the high variation in the scores. An example of this issue is the design, which is considered as the least important attribute of the store choice and posses the highest deviation from the mean.

*Chapter 5*

# MODEL ESTIMATION AND RESULTS

## 5.1 Factor estimation

On the next stage of the study the obtained results are evaluated through the factor analysis. Within the first part of the survey the consumers were asked to rank the main attributes of the supermarket. There were 7 dimensions to be ranked that represent the main part of the survey. In the Table 2 below are presented main results for the 18 questions that consumers were asked, the questions are grouped into 7 categories. These seven factors represent 78 percent of the variance of the 21 variables. This percentage is adequate since the required percentage in the social sciences is 60%. The significance of the factor loading is presented in the Table 2.

The factor analysis was performed using Stata software package standard procedure. The method that was used is the Principal Factor Analysis. The choice of seven representative factors was confirmed by the application of the maximum-likelihood factor method in the factor analysis to the data.

Table 2. Factors for the survey.

| Name | Description | Mean | Standard Deviation | Factor loading |
|---|---|---|---|---|
| **Convenient for the transport** | | | | |
| tr_publ | Easy access by Public transport | 7.47 | 2.27 | **0.69** |
| tr_parc | Convenient parking | 5.82 | 3.29 | 0.47 |
| **Checkout and shopping assistance** | | | | |
| serv_speed | Fast checkout | 5.84 | 2.43 | **0.89** |
| serv_ass | Shopping assistants are courteous | 6.21 | 2.59 | 0.26 |
| serv_qual | Qualified assistance | 5.21 | 2.43 | 0.38 |
| **Supermarket design and physical facilities** | | | | |
| des_crowd | Not crowded store | 5.61 | 2.38 | **0.69** |
| des_move | Easy to move around the store | 6.34 | 2.56 | 0.49 |
| des_find | Easy to find products | 6.55 | 2.31 | 0.41 |
| **Quality and Range of products** | | | | |
| prod_basis | Store has basic products | 7.53 | 1.97 | 0.14 |
| prod_varie | Good variety of products | 7.00 | 2.18 | 0.21 |
| prod_fresh | Fresh products | 6.79 | 2.09 | 0.16 |
| prod_qual | High quality products | 7.24 | 1.50 | **0.24** |
| **Pricing Policy** | | | | |
| price_low | Low price store | 6.08 | 2.29 | **0.77** |
| price_prom | Promotional offers | 5.58 | 2.68 | 0.2 |
| price_adv | Advertising of products | 5.63 | 2.35 | 0 |
| **Opening hours** | | | | |
| op_late | Open late at night | 6.03 | 3.39 | 0.13 |
| op_early | Open early in the morning | 6.50 | 3.13 | **0.14** |
| **Location** | | | | |
| loc | Well located | 7.47 | 2.29 | **0.36** |

On the first step I assume 7 factors among the variables of the survey. Based on the estimation results, I was not able to reject the null hypothesis of no more than 7 factors in the data. The resulting $\chi$ squared statistics for 48 degrees of freedom is 46.12 with the probability the statistics to exceed the critical value that is equal to 0.5502. On the second step I performed the estimation with the assumption of

6 factors in the model. The estimation showed the possibility of rejection of the null hypothesis of no more that 6 factors in the data with 10 percent level of confidence. The resulting $\chi$ squared statistics for 46 degrees of freedom is 76.71 with the probability the statistics to exceed the critical value that is equal to 0.0717. The factor specification has also confirmed the initial guess of no more than 7 factors in the model.

The final step of the Factor analysis is the choice of the surrogate variables for each of the factors. The surrogate variables were chosen based on the highest factor loading of the variable within the cluster. Those surrogate variables will represent the factors in the further regression analysis. The surrogate variables are presented in the Table 3.

Table 3. Surrogate variables for survey.

| | | |
|---|---|---|
| Factor 1 | tr_publ | Easy access by Public transport |
| Factor 2 | serv_speed | Fast checkout |
| Factor 3 | des_crowd | Not crowded store |
| Factor 4 | prod_qual | High quality products |
| Factor 5 | price_low | Low price store |
| Factor 6 | op_early | Open early in the morning |
| Factor 7 | loc | Well located |

The surrogate variable for the transport convenience is accessibility by public transport. We can compare result obtained to the results of Colomé and Serra (2003). For the Spanish survey authors obtained the highest factor loading for the variable that represents the convenience for parking. This result is intuitive as the population in the considered area has quite low standard of living and as a result public transport is the primary way of transportation. On the contrary, for the Spanish consumers, the car is the major way of transportation due to the high living standard in Europe.

The choice of surrogate variables for the shopping service is the speed of service, which also represents the common sense, as the fast checkout is the primal consideration in the store services. The results in the paper of Colomé and Serra (2003) coincide with the results for Ukraine. The highest factor loading in this group for the Spanish survey was also factor representing fast checkout.

The surrogate variable representing the design and physical facilities is variable that characterize the crowd in a store. We can compare it to the result of Colomé and Serra (2003), where the choice of the surrogate variable for the given cluster is easiness to move around the store. The difference in the factor selection could be explained by the distinction in the consumer's behaviour in the countries of interest.

The surrogate variable for the group that corresponds to the product characteristics is product quality. This result corresponds to the result for the Spanish survey, which suggests that the choice of surrogate variable confirms the common belief of the product quality for the consumers.

As the surrogate variable for the supermarket price policy factor analysis pointed low price image of a store, which confirmed the hypothesis of the highest relevance of the low price for the consumers in the region under the study.

## 5.2 Model calibration

The surrogate variables found using the factor analysis are now can be incorporated into the MCI model. Those variables will represent the key supermarket attributes $A_{kij}$. As a result I obtain the initial specification of the model (3.3):

$$p_{ij} = \frac{tr\_publ_{ij}^{\beta 1} * serv\_speed_{ij}^{\beta 2} * des\_crowd_{ij}^{\beta 3} * prod\_qual_{ij}^{\beta 4} *}{\sum_{j=1}^{m}(tr\_publ_{ij}^{\beta 1} * serv\_speed_{ij}^{\beta 2} * des\_crowd_{ij}^{\beta 3} * prod\_qual_{ij}^{\beta 4} *}$$

$$\frac{* low\_price_{ij}^{\beta 5} * op\_early\_speed_{ij}^{\beta 6} * loc_{ij}^{\beta 7}}{* low\_price_{ij}^{\beta 5} * op\_early\_speed_{ij}^{\beta 6} * loc_{ij}^{\beta 7})} \quad (4.1)$$

In the presented model the only parameters that remains unknown are the sensitivity parameters $\beta_k$. I obtain these parameters through the calibration of the model based on the initial specification. The calibration was performed in Stata 6.0 software package using the regression techniques, namely OLS with HAC Consistent Covariances (Newey-West) and ML regression model.

On the first step of calibration, I created consumer zones from individual consumer responses about the location of the supermarket. On the second step I computed probability that consumer from zone *i* will shop in store j. The probability is computed using the frequency method, where the probability of consumer from zone *i* to go to store j is computed as the number of consumers from zone *i* that were questioned in store j devised by total number of consumers in zone *i*. On the final step I perform the regression on the log-centred equation (4.2).

$$\ln\left(\frac{p_{ij}}{\hat{p}_i}\right) = \beta_1 \cdot \ln\left(\frac{tr\_publ_{ij}}{\overline{tr\_publ_{ij}}}\right) + \beta_2 \cdot \ln\left(\frac{serv\_speed_{ij}}{\overline{serv\_speed_{ij}}}\right) + \beta_3 \cdot \ln\left(\frac{des\_crowd_{ij}}{\overline{des\_crowd_{ij}}}\right) \quad (4.2)$$

$$+ \beta_4 \cdot \ln\left(\frac{prod\_qual_{ij}}{\overline{prod\_qual}}\right) + \beta_5 \cdot \ln\left(\frac{price\_low_{ij}}{\overline{price\_low_{ij}}}\right) + \beta_3 \cdot \ln\left(\frac{open\_early_{ij}}{\overline{open\_early_{ij}}}\right)$$

$$+ \beta_7 \cdot \ln\left(\frac{loc_{ij}}{\overline{loc}}\right) + \beta_8 \cdot st + \beta_9 \cdot near + \ln\left(\frac{\zeta^*_{ij}}{\hat{\zeta}^*_i}\right)$$

The regression summary is presented in the Table 4. The R-squared of the OLS model is equal 21%. This result is not very high but it is quite reasonable given that the data under consideration is cross-sectional data. The interpretation of the regression results could be valid if there is no specification error in the model. The verification of the model validity was done using standard tests. I have performed the Ramsey Reset test to identify the following specification errors:

- Omitted variables, to test for the situation when the model does not include all relevant variables.
- Incorrect functional form, it can be the case when some or all of the variables in the regression should be transformed to logarithm, raised into power or transformed in the other way.
- Correlation between independent variables and error term that may be caused by measurement error.

F-statistics for the Ramsey Reset test is 1.24. The probability of F-statistics is rather high, 0.27. As a result we cannot reject the null hypothesis about the specification error at 10% level of significance and conclude that there is no specification error.

To test wether the disturbances are normally distributed I have used Jarque-Bera test. The resulting Jarque-Bera statistics is 4.8, which imply that we cannot reject the null hypothesis of the normally distributed disturbances at the 5 % level of confidence and we can conclude that the disturbances are distributed normally.

Another relevant test for the residuals is the test for autocorrelation. The primary test here is Durbin-Watson test. The DW-statistic is a test for first-order serial correlation. If the DW-statistics turns to be lower than 2, it points to a positive autocorrelation. The result of DW statistics in my regression is 0.34, which is a sign of strong positive autocorrelation. The same conclusion can be made from

the inspection of the Correlograms and Q-statistics of the data. The Correlograms shows strong evidence of a positive autocorrelation.

The model was also tested for the heteroskedasticity applying the standard White-heteroskedasticity test. The p-value of F-statistics in the considered model is 0.07, which is not very high. However, it exceeds the critical 0.05. Nevertheless, the p-value does not exceed the critical 0.1. Consequently, we cannot reject the null hypothesis of heteroskedasticity at 5% level of confidence. However, we can reject the null hypothesis and at 10% level of confidence. Therefore we can conclude that there is a rather high change of the heteroskedasticity presence in the model.

Based on the tests results we can conclude that there is a high chance of positive autocorrelation and heteroskedasticity in the data. As a result the coefficients estimates are unbiased and consistent but they are not efficient. For that reason I use OLS with HAC Consistent Covariances (Newey-West) and ML regression models in the MCI calibration.

Table 4. Regression results.

| Variable | Coefficient OLS | Coefficient GLM | Std. Error | Std. Error | p-value OLS | p-value GLM |
|---|---|---|---|---|---|---|
| **TR_PUBL** | **0.094*** | **0.094**** | **0.052** | **0.459** | **0.075** | **0.041** |
| SERV_SPEED | 0.011 | 0.011 | 0.033 | 0.032 | 0.738 | 0.727 |
| DES_CROWD | -0.001 | -0.001 | 0.039 | 0.034 | 0.972 | 0.968 |
| PROD_QUAL | -0.065 | -0.065 | 0.073 | 0.069 | 0.372 | 0.348 |
| OP_EARLY | -0.034 | -0.034 | 0.037 | 0.035 | 0.356 | 0.335 |
| LOC | -0.077 | -0.077 | 0.052 | 0.050 | 0.143 | 0.126 |
| **ST** | **0.162*** | **0.162***** | 0.083 | 0.051 | **0.053** | **0.001** |
| **PRICE_LOW** | **0.102***** | **0.102***** | 0.041 | 0.040 | **0.013** | **0.011** |
| **NEAR** | 0.057 | **0.057*** | 0.066 | 0.033 | **0.388** | **0.092** |
| C | -0.786 | -0.786 | 0.067 | 0.036 | 0.000 | 0.000 |

By inspection of the estimation output for both the regressions we can conclude that OLS and ML provide a reasonably close results.

As can be seen from the table, the significant variables for the both regressions are tr_publ, loc, price_low and st. The ML estimation also points a dummy variable near as significant for our study. The sensitivity parameters $\beta_{k'}$ are the regression coefficients for the surrogate variables.

The highest sensitivity parameter among the key supermarket attributes is the parameter for variable price_low that reflects the consumer valuation for the price policy of a store. This result is rather intuitive, as the price policy of a store remains the most significant parameter of the supermarket choice for the Ukrainian consumers. This conclusion is also confirmed by the survey results. When consumers were asked about the importance of different supermarket attributes for the store choice they have given the priority to the low price image. This result was presented in Table 1. The positive sign for the parameter for the variable price_low indicates that a supermarket with higher score for the price image would have higher probability of being shopped in.

Another key supermarket attribute that is significant at 1% level of confidence is represented by surrogate variable tr_publ. This attribute reveals attitude of the consumers towards convenience for transportation in the supermarket. This result is also intuitive as consumers usually do shopping both in weekends and during the working week. For the latest type of hopping the factor of transport convenience is highly relevant since shoppers usually go to the store after the work. The estimation of this sensitivity parameter is also positive which is expected sigh since the more convenient is store for the transportation the higher will be probability that the consumer will those this very supermarket.

The other key supermarket attributes turn to be insignificant which indicated that we do not need to include them into the final version of the model. This result is rather surprising since such characteristics of the store as quality of products, service and location are quite important for the consumer choice. The

significance was confirmed by the high importance scores that consumer have given to this characteristics in the survey, which can be seen from Table 1. Such attribute as product quality and range has almost the same score as the low price image (4.22 versus 4.2). However, this phenomenon has a reasonable explanation. The supermarkets under the study could be very similar in product characteristics. As a result, the consumers will treat these stores as equal for this very attribute and will made the choice of the supermarket based on other characteristics. The same logic is applicable to the other key supermarket attributes, which turn to be insignificant.

We can compare the result obtained for Ukrainian market with the result for Spanish and British markets, which is taken from Colomé and Serra (2003) paper. The conclusions for the Spanish survey are identical to the results obtained for the case of Ukraine. The significant surrogate variables for the Spanish case are also those representing the cluster of price image of the store and the cluster representing the convenience for transportation. However, in the Spanish survey transport convenience cluster was characterized by the parking accessibility, rather than public transportation.

The results for the British survey also point to a price policy of a store as a significant factor for the consumer choice. The other significant factor for the consumer choice in Great Britain is variable that corresponds to the quality and range of product cluster. This result could also be explained by the differences in the local market.

An important conclusion from the comparison is that in all three cases the price policy of a store turn to be significant factor in the consumer choice, which is confirmed both by the general intuition and the survey results.

Based on the regressions results we are also able to get the estimated sensitivity parameters $\beta_k$, which now could be plugged in the model. As a result I obtain the final version of the log-centred estimated MCI model.

$$\ln\left(\frac{p_{ij}}{\hat{p}_i}\right) = 0.09 \cdot \ln\left(\frac{tr\_publ_{ij}}{\overline{tr\_publ_{ij}}}\right) + 0.1 \cdot \ln\left(\frac{price\_low_{ij}}{\overline{price\_low_{ij}}}\right) \quad (4.3)$$

On the next step I transform the model into the original representation. The final version of the estimated MCI model is as following.

$$p_{ij} = \frac{tr\_publ_{ij}^{0.09} * price\_low_{ij}^{0.1}}{\sum_{j=1}^{m}(tr\_publ_{ij}^{0.09} * price\_low_{ij}^{0.1})}, \text{ where} \quad (4.4)$$

$p_{ij}$ – The probability that consumers located at $i$ will make purchases in store $j$.

*tr_publ* – valuation by zone $i$ consumers of the convenience for the transportation of the store j.

*price_low* – valuation by zone $i$ consumers of the low price image of the store j.

## 5.3 Application of the results

An application of this research is construction of a revised version of the maximum capture problem for the supermarket sector that will incorporate the revealed preference of consumer in the store-choice behaviour. The formal representation of the revised model is presented below.

$$MAXZ = \sum_{i \in I} \sum_{j \in J} a_i \rho_{ij} x_{ij} \text{ , subject to} \quad (4.5)$$

$$\sum_{j \in J} x_{ij} = p + q, \forall i \in I$$

$$x_{ij} \le x_{jj}, \forall i \in I, \forall j \in J$$

$$\sum_{j \in J} x_{jj} = p$$

$$x_{ij} = \{0,1\}, x_{jj} = \{0,1\}, \forall i \in I, \forall j \in J$$

$$p_{ij} = \frac{tr\_publ_{ij}^{0.09} * price\_low_{ij}^{0.1}}{\sum_{j=1}^{m} (tr\_publ_{ij}^{0.09} * price\_low_{ij}^{0.1})}$$

In the revised model the possibility of several stores to be assigned to one consumer zone is added. This will better represent the true situation since in reality the consumers from one consumer zone do shop in different supermarket. In the revised model the consumer can shop in p new and q existing supermarkets. The parameter $p_{ij}$ is determined using the results of the calibration of the Multiplicative Competitive iteration Model. As a result the revised version of the model will incorporate the different store attributes into the model of supermarket choice. The parameter $p_{ij}$ is rather sensitive to the local situation since the consumer behaviour is highly subjective. This paper represents the calibration of the model parameters for the local Ukrainian market.

*Chapter 6*

# CONCLUSIONS

The purpose of this paper is application of the methodology for deciding which store attributes should be included into the maximal capture model to the local situation of Ukrainian retail market.

The research was performed in the course of four major steps. On the first step I designed a survey for collecting of the information that reveals the consumer preference in a given area. During the second step I have carried out the analysis of the survey data. The technique employed for the study was factor analysis. On this stage of the research I have chosen the surrogate variables to characterize the groups of key supermarket attributes. The surrogate variables for the store attributes represent convenience for public transportation, speed of checkout, crowdedness of a store, quality of product, price policy, hours of opening and the location of the supermarket. The resulting surrogate variables were compared to the results obtained by Colomé and Serra (2003) for both Spanish and British surveys.

On the third step the surrogate variables obtained on the previous stage were incorporated in to the MCI model. The final step was calibration of the MCI model using the regression techniques. Based on the output of estimation I have chosen two supermarket attributes that are significant for the study, which are the convenience for transportation and price policy of a store. The significant attributes were also compared to the results of Spanish and British survey. The resulting attributes for the Spanish survey coincide with the Ukrainian sample result. The results for the British survey have only partial correspondence, since

in British study the variable for transport accessibility turn to be insignificant whereas the variable for the quality of products show the significance for the study. The regression also provides the estimation of the sensitivity parameters in the MCI model. With the estimated sensitivity parameters model got final specification. As a consequence I obtained a model for the probability of consumer from a specified zone to shop in a particular supermarket. The resulting probability was further incorporated into the MaxCap model, which characterize the modification of the model that encounter for the other supermarket attributes other than location. Therefore the resulting model is more applicable to the real world environment and provides retail firm with powerful tool for decision upon the stores location.

Although the paper represents a powerful model for the deciding on the optimal store location choice, the methodology contains some limitations. However, those limitations indicate the directions to further research in the area.

The first limitation connected with the MCI model assumptions. The model itself supposes that consumers have compensatory utility functions. However, consumers eliminate from the consideration the stores, which do not meet the specified minimum requirements level. In our case the limitation is not valid since the supermarkets under consideration do cover the minimum requirement set.

Another fact that limits the analysis is similarity of the supermarkets under the study. This may result in low significance for the MCI model of the attributes that are close to each other. As a consequence, the attributes that in reality are significant for the consumer choice may turn to be insignificant for the current study. However, expanding of the set of stores under the study will eliminate this problem, especially if we include into the analysis the stores that significantly differ. Unfortunately, such an expanding was not possible in the current study

due to the operational and financial constraints. Therefore it provides us with a direction for the future research.

Another constraint is the determinations of the probability of a consumer from zone *i* to shop in a store j. In the present research the probability is determined by the simple frequency method. However, more applicable technique here would be the estimation of such probability for each consumer. In current study this information was not collected due to the operational difficulties.

The present research represents a resolution of the updated version of the maximum capture problem. The further direction of a research would be the solution of the MaxCap problem for the retail chain entering the spatial market. The resolution of the model may involve the process of computation within the simulated network for verification of the optimality of the obtained solution. After the optimal solution on the simulated network would be verified, model would be ready for application to the real world. The resulting model would provide the firm that is going to enter the spatial market with the model of optimal supermarket location that takes into account the different attributes of the supermarket attractiveness.

# BIBLIOGRAPHY


M. Birkin, G.Clarke and lee *Douglas Optimizing spatial mergers: commercial and regulatory perspectives* Progress in Planning #58, 2002, 229-318

O. Berman, D.Krass *The generalized maximal covering location problem* Computers & Operations Research 29, 2002, 563–581

K.Campo, E. Gijsbrechts *Should retailers adjust their micro-marketing strategies to type of outlet? An application to location-based store space allocation in limited and full-service grocery stores*, Journal of Retailing and Consumer Services 11, 2004, 369–383

R.L. Church, C. ReVelle *The maximal covering location problem* Papers of Regional Science Association, 32, 1974, 101–18.

G. Clarke *Changing methods of location planning for retail companies*, GeoJournal 45**:** 1998, 289–298.

R.M. Clarkson *UK supermarket location assessmentof location planning for retail companies*, International Journal of Retail and Distribution Management 24, 1996, 22-33

R. Colomé, D. Serra *Supermarket Key Attributes and Location Decisions: A Comparative Study between British and Spanish Consumers* Working Paper, October 2003

C.S. Craig, A. Ghost, S. McLafferty *Models of Retail Location Process*, Journal of Retailing, v60, 1984, 5-36

C.S. Craig, A. Ghost *Fermulating Retail Location Strategy in a Changing Environment*, Journal of Marketing, v47, 1983, 56-68

R.L. Davies *Store Location and Store Assessment Research*, Wiley, Chichester, 1977

D. Gautschi *Specification of Patronage Models for retail center Choice*, Journal of marketing research, 18, 1981, 162-174

H. Hotelling *Stability in Competition*, Economics Journal, v39, 1929, 41-57

D. Huff *Defending and Estimating a Trading Area*, Journal of Marketing 28, 1964, 34-38

O. Karasakal, E.. Karasakal *A Maximal Covering Location Model in the Presence of Partial Coverage* Computers & Operations Research 31, 2004 1515–1526

V. Kumar, Kiran Karande *The Effect of Retail Store Environment on Retailer Performance* Journal of Business Research 49, 2000, 167–181

J. Louviere, G. Woodworth *Design and Analysis of Simulated Consumer Choice of Allocation Experiments: an Approach based on Aggregated Data* Journal of marketing research, 20, 1983, 350-367

J. Louviere *Hierarchical Information Integration: A New method for the Design and Analysis of Complex Multiattribute Judgements Problems* Advances in Consumer Research, Proto Utah: Association for Consumer research, 1984, 148-155

R. Murto, The Nordic Research Network on Modelling Transport Land-Use and the Environment *The Optimal Location of Grocery Stores - case transport* Fifth workshop, Nynäshamn September 28-30, 2001

M Nakanishi, M. Cooper *Parameter estimation for a Multiplicative Competitive Interaction Model - Least Square Approach* Journal of Marketing vol 11, 1974, 303-311

W.J. Relly, *The Law of Retail Gravitation*, New York, Knickerbocker Press, 1929

L.P. Simkin *Evaluating a store location*, International Journal of Retail and Distribution Management 18, 1990, 33-38

C.C. Yrigoyen, J.V. Otero *Spatial interaction models applied to the design of retail trade areas*, 1998 European Regional Science Association conference papers.

APPENDIX A

SURVEY

**1) Which importance scores would you give to the following supermarket attributes?**

| Not important | Rather unimportant | Middle importance | Rather important | Very important |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

1. Location 1 : 2 : 3 : 4 : 5 :
2 Low prices 1 : 2 : 3 : 4 : 5 :
3. Quality and range of products 1 : 2 : 3 : 4 : 5 :
4. Hours of opening 1 : 2 : 3 : 4 : 5 :
5. Speed of service 1 : 2 : 3 : 4 : 5 :
6. Quality of service 1 : 2 : 3 : 4 : 5 :
7. Design 1 : 2 : 3 : 4 : 5 :

**2) How would you evaluate the following attributes of this supermarket?**

**Convenient for the transport**

- ✓ Easy access by Public transport
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Convenient parking
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**Checkout and shopping assistance**

- ✓ Fast checkout
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Shopping assistants are courteous
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Qualified assistance
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**Supermarket design and physical facilities**

- ✓ Not crowded store
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Easy to move around the store
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Easy to find products
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**Quality and Range of products**

- ✓ Store has basic products
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Good variety of products
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Fresh products
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ High quality products
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**Pricing Policy**

- ✓ Low price store
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Promotional offers
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Advertising of products
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**Opening hours**

- ✓ Open late at night
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

- ✓ Open early in the morning
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**Location**

- ✓ Well located
  1 : 2 : 3 : 4 : 5 : 6 : 7 : 8 : 9 : 10 :

**3) How close is the supermarket to your home?**

less than 5 minutes -----5-10 minutes------10-15 minutes ---15-20 minutes----more than 20