A discusión

LOCATION STRATEGIES BASED ON DISCRETE CHOICE MODELS: AN EMPIRICAL APPLICATION TO SUPERMARKET LOCATION *

I. Arribas, J.L. Cervera and J.E. Vila **

WP-AD 2004-26

Corresponding author: Ivan Arribas, University of Valencia, Departamento de Análisis Económico. Edificio Departamental Oriental. Campus dels Tarongers; Avda. dels Tarongers, s/n. 46022 Valencia (Spain).
Tel.: 34-96-3828246, Fax number: 34-96-3828249; e-mail: ivan.arribas@uv.es

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

Primera Edición Junio 2004.

Depósito Legal: V-3006-2004

* We wish to thank financial aid from the Valencian Institute of Economic Research (IVIE).
** I. Arribas: Universitat de València, Departamento de Análisis Económico, Edificio Departamental Oriental. Campus dels Tarongers; Avda. dels Tarongers, s/n. 46022 Valencia (Spain). Teléfono: 34-96-3828246, Fax: 34-96-3828249; e-mail: ivan.arribas@uv.es; J.L. Cervera: Devstat; J.E. Vila: Universitat de València, Departamento de Análisis Económico. Edificio Departamental Oriental, Campus dels Tarongers, Avda. dels Tarongers, s/n. 46022 Valencia (Spain), Teléfono: 34-96-3828246, Fax: 34-96-3828249; e-mail: jose.e.vila@uv.es

LOCATION STRATEGIES BASED ON DISCRETE CHOICE MODELS: AN EMPIRICAL APPLICATION TO SUPERMARKET LOCATION

I. Arribas, J.L. Cervera, and J.E. Vila

ABSTRACT

In this paper we present a theoretic two-stage model for retailers location and consumers purchase decision. Retailers decision problem is formalized in terms of a zero-sum game, whose payoffs refers to retailers' market share and consumers decision problem is formalized in terms of a discrete choice model, based on random utilities. The theoretical models provide forecasting of equilibrium market shares and the locations to be chosen by retailers, in terms of the geographic distribution of the underlying location space (constituencies of the town), population distribution and characteristics (types) of the consumers.

The theoretic model is applied to analyze empirical supermarket distributions on three villages of the Region of Valencia: Requena, Segorbe and Utiel. Departing from actual market shares in these three towns, a function relating types (young and older consumers) and sales is estimated. This estimate allows to calculate the payoff matrices for the location games corresponding to the three towns and to find out that there is just one equilibrium in pure strategies for each of them. A comparison of these equilibria and the actual location patterns of supermarkets, shows that our model explain quite well decision location but it can be improved to obtain more precise forecasts of actual market shares.

Keywords: Hotelling, Industrial Organization, Choice Model

1 Introduction

Much attention has been paid by economists to location problems since the early work of Hotelling (1929). In this framework, a wide series of interesting stylised models has been developed, which integrate location decisions with decision on other strategic variables such as price, capacity, etc[1]. These models generalize Hotelling's seminal idea by introducing (1) a more sophisticated 'geography' (firms do not choose locations in a segment but in circumferences, areas of a plane, etc., (2) non-linear transport cost functions, (3) alternative cost functions for firms or (4) alternative decision variables, such as capacity.

In a first moment, location theory appeared as an instance of differentiation theory to relax price competition and to break the Bertrand 'paradox', i.e. the idea that two firms are enough for competitive outcomes. However, the practical implications of these problems in many areas of management and policy design made location problems interesting by themselves, both from theoretical and empirical viewpoints. Roughly speaking, geographic factors do influence both firms' and consumers' behavior. A firm choosing a specific location creates a partial product differentiation, and generates a surrounding influence area including neighboring households that choose to buy the firms' product.

The literature describes two main alternative approaches for modelling the process that establishes influence areas (Chasco, 1997): descriptive-deterministic and explanatory-stochastic models. While the former do not intend to explain the factors that create influence areas, but to fit observed data by ad hoc mathematical models, the latter are based on hypotheses on consumers' rational behavior, which decide the geographic market where they consume after a utility maximisation process (generally, the minimisation of commuting costs) given the location of the establishments. Explanatory-stochastic include Spatial Interaction models (Huff, 1963), discrete choice models (McFadden,1977; Fotheringham, 1989), and models of Direct Utility Evaluation (Louvière, 1983).

The main goal of this paper is to establish a bridge between the theoretical and empirical approaches. We depart from of theoretical models of both consumption and localization decisions, and then we generate some simple

[1]For a general survey of the development on the theory of 'Hotelling's models' see, for instance Brenner (2001), Harter (1996), Gabszwewicz and Thisse (1992), Lancaster (1990) , Waterson (1989) and Osborne and Pitchick (1987), among others.

forecasts to be empirically compared with the situation in real markets. Our theoretical model considers simultaneously the interactions among firms and consumers. Hence, the problem of identifying influence areas is modeled as a two-stage decision model where first firms choose their locations based on their previous and private knowledge of the distribution of the consumers' population and tastes, with establishment costs high enough to discourage firms from changing their locations and in the second stage consumers choose the establishments where to buy. We refer to the first stage as the 'location game' and to the second one as the 'consumer choice model'.

To reach that goal, we need to establish some assumptions on:

- Geography of locations
- Competition models of retailers
- Choice models of consumers

1.1 The geography of locations

The spatial geography in which the location game takes place is a fundamental issue. The need of empirical testing forces to define a 'geography' of the underlying spatial framework that is realistic enough to allow contrastable forecasting and adapted to available data. The need of realism implies to consider location in the plane, instead of the usual one-dimension space more commonly considered in the literature. The ideal underlying space is, then, a street and road map of the analyzed towns. In this case, a consumer can be characterized by a pair of real numbers, corresponding to the geographic coordinates of the house where she is living. On the other hand, firms choose a pair of real numbers within the area, representing again the geographic coordinates of the selected locations, and establish their new premises there. Although appealing, this model will not be considered, the reason being that its information requirements to produce forecasting can not be satisfied by secondary statistical sources: to deal with this model we would need to know characteristics and preferences of a sample of individual consumers together with their exact co-ordinates. Since we are constricted to the public secondary data, the smallest geographical unit for which relevant information is available are constituencies[2] and, even on this aggregation level, the only easily available data concern the number of inhabitants by age, sex, education and nationality. In this paper, the geographical space is discretized, consisting in a finite selection of locations (the constituencies). For our purposes, we shall assume that the competing firms choice the establishments' locations at the constituency level.

One of the most critical points in localization models is the definition of the transport cost function. In the literature see (for instance, Brenner, 2001) transport costs are defined as a deterministic (usually linear) function on the Euclidean distance between the consumer and the firms. However, a realistic model should consider the topology induced by the existence of accidents (natural or man-made) such as rivers or roads that invalid the Euclidean formula for its use in the calculation of distances. Different metrics such as the 'Manhattan' or block distance have been proposed for urban models. Alternatively, these metrics can be summarized by the use of time instead of spatial distance. This approach seems appealing for empirical applications, since consumers are mainly concerned not on the spatial distance to reach a specific location but on the time they need to spend to get there. Then, in our model of real localization problem, we define a distance between two constituencies based on the average time that is needed to commute between one and the other. Specifically, we define the distance function as the time needed to go to the geographical center of a constituency to the geographical center of another. We also assume that the distance from a constituency to itself is not null and it depends on the time that consumers spend to go from its center to its border.

1.2 Competition models of retailers

As a difference with the standard Hotelling approach, our model does not consider an intermediate phase where, once localizations are chosen, firms decide prices and communicate them to their potential buyers. We assume, then, that prices for both firms are identical and exogenous to the problem. Even if this hypothesis may look artificial and unrealistic, price exogenity is one of the most common frameworks for practical in-depth location decisions:

[2]By constituencies we mean "seciones censales", the minimun divison used for statistical purposes in Spain.

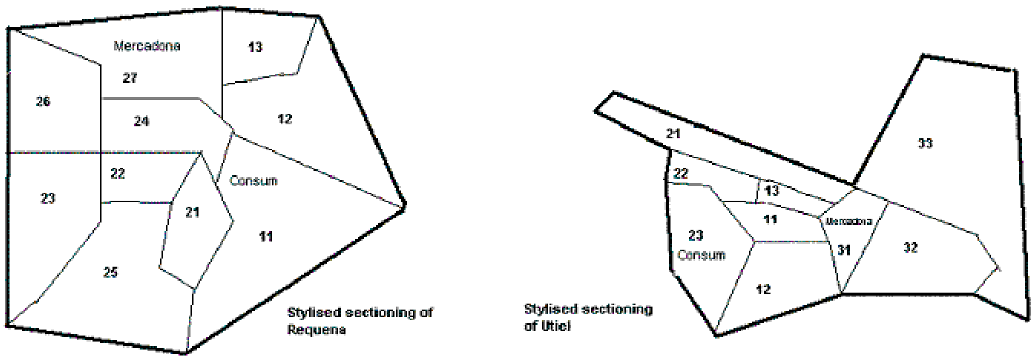

Figure 1: Stylised representation of the consituencies in Requena (left) and Utiel (right) and supermarket locations.

those made by the expansion departments of large retailing and restauration companies. These companies set prices based on global competition strategies that, most often, are close to those of their main competitors. Clear examples of these situations, are Burger King / McDonald's prices, or, in a more local framework, Mercadona / Consum[3] average prices. These uniform prices are applied with no exception to all of their current premises and potential openings. Then, when two of these brands are considering to inaugurate new premises in a town, our model may provide a good description of their decision making.

1.3 Choice models of consumers

Following an explanatory-stochastic approach, we assume that consumers' behavior can be describe by a Discrete Probability Choice Model (McFadden, 1977). Formally, a probabilistic choice system (PCS) is defined as $(I, Z, \xi, C, \Theta, P)$ where I is the set of indices for the alternatives, Z the universe of measured attributes of alternatives, $\xi : I \longrightarrow Z$ a mapping specifying the observed attributes of the alternatives, C a family of finite, non-empty choice sets from I, Θ the universe of vectors of measured characteristics of

[3] Mercadona and Consum are two Spanish supermarket firms that are sales leader in the Region of Valencia.

individuals making choices (the consumers), and $p : I \times 2^C \times \Theta \longrightarrow [0,1]$ the choice probability function, such that $p(i|B,\theta)$ is the probability of alternative i being selected given that the selection must be made from the choice set $B \subset C$ and that the decision maker has characteristics $\theta \in \Theta$.

For the case analysed, I is the set of indices for market places (establishments or firms locations), each of them with attributes Z that may include the spatial coordinates, the selling price, amenities offered, and other. Θ may specify demographic or economic variables of the consumers, or any other aspect influencing tastes.

The distribution of tastes in the population of decision-makers (consumers) is given by a probability measure $\mu(.|\theta)$ in the space $\mathfrak{U}(I)$ of utility functions with arguments in I, depending on their characteristics θ.

The introduction of a supplementary random component in the utility function leads to the Random Utility Maximisation (RUM) paradigm, extensively studied again by McFadden (1977), which allows considering a population of consumers with both known and unmeasured covariates influencing their decision, and their distribution in a geographical space. The formal integration of all information elements may be provided by utility functions of the form

$$U \equiv W + \varepsilon$$

where W is the deterministic or systematic part of the utility and ε is a random term, capturing the uncertainty whose sources are the unobserved attributes of the alternative establishments, the unobserved individual characteristics (such as psychological factors), measurement errors (for example, of distances and transportation costs), and other.

MacFadden demonstrates that a PCS is compatible with the RUM hypothesis (or can be generated from the RUM hypothesis) and a family of choice sets $B \in \mathfrak{B}$ via the following mapping: $p : I \times 2^C \times \Theta \longrightarrow [0,1]$ defined by

$$p(i_k|B,\theta) = \mu\left(\{U \in \mathfrak{U}(I) \quad / \quad U(i_k) = \max_j U(i_j)\}, \theta\right)$$

for each $B = \{i_1, ..., i_n\} \in C$,and $\theta \in \Theta$.

Finding econometrically feasible PCS consistent with RUM is done then by generating choice probabilities p from parametric families of probabilities

μ. Specifically, following Luce's Utility Axiom (Luce, 1959) we assume that[4] the probability to buy in retailer located in i_k is given by the ratio of the utility provided by the purchase in retailer placed in i_k over the sum of utilities of buying in all the retailers. Formally,

$$p(i_k|B,\theta) = \frac{U(i_k|B,\theta)}{\sum\limits_{i\in B} U(i|B,\theta)}.$$

This model allows for introduction, in the evaluation of the utilities, of explanatory variables for the individual decision, as well as interaction among establishments.

The structure of the rest of the paper is as follows. Section 2 shows a very simple example of the formal model, which is established in Section 3. Section 4 is devoted to solve the model and to deduce the outcomes to be empirically tested. The empirical cases to be analyzed are presented in Section 5 and the correspondence between the model outcome and the empirical data is checked in Section 6. Section 7 summarizes the main results of the paper and points out further research on these topics.

2 Example

Let us assume that two different brands, say i and j, want to open a new supermarkets in a town. Supermarkets differ only in the trade mark and consumers do not have preferences for any of these brands. The town has two separate constituencies to be called c_1 and c_2. There are two different types of consumers in the town. Consumer type θ_1 cannot drive and only buy in the area where they are living, while consumers of type θ_2 can drive to any point of the town. Assume that the probability of a consumer of type θ_2 to buy in a given supermarket is proportional to the inverse of the time they need to arrive from home to the location of that supermarket. There are n_{sk} consumer of type θ_s in constituency c_k, $k = 1, 2; s = 1, 2$, and this information is common knowledge to retailers and consumers. Both types of consumers buy for the same amount of money, say 1 currency unit to normalize. Finally, let us assume that the average time spent to drive to a

[4]Huff (1963) proposed a similar consumer's choice probability, but without considering different types of consumers.

supermarket placed in the same area where a consumer lives is T, meanwhile the driving time to commute to a different area is aT, where $a > 1$.

Let $p(c_l|\theta_i, c_k)$ denote the probability to buy in constituency c_l of a consumer of type θ_i who lives in constituency c_k. From the assumptions, it is easy to check that:

$$\begin{aligned} p(c_l|\theta_1, c_k) &= \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \\ p(c_l|\theta_2, c_k) &= \begin{cases} \frac{a}{a+1} & k = l \\ \frac{1}{a+1} & k \neq l \end{cases} \end{aligned}$$

Then, the expected number of consumers who will buy in region c_k are given by

$$\pi(c_k) = n_{1k} + \frac{a}{a+1} n_{2k} + \frac{1}{a+1} n_{2l}$$

$k = 1, 2, l = 1, 2, k \neq l$.

Supermarkets' problem to choose specific locations to open their new premises, may be modelled then as zero-sum game. Total sales of a supermarket depend then on its own location and on the location of its competitor. Let us denote by $\pi_i(c_l, c_k)$ the total sale of company $i = 1, 2$ when supermarket 1 is placed in region c_l and supermarket 2 is placed in c_k. We have then,

$$\begin{aligned} \pi_1(c_l, c_k) &= \begin{cases} \frac{1}{2} n_{1l} + \frac{1}{2}(n_{21} + n_{22}) & k = l \\ n_{1l} + \frac{a}{a+1} n_{2l} + \frac{1}{a+1} n_{2k} & k \neq l \end{cases} \\ \pi_2(c_l, c_k) &= \begin{cases} \frac{1}{2} n_{1l} + \frac{1}{2}(n_{21} + n_{22}) & k = l \\ n_{1k} + \frac{a}{a+1} n_{2k} + \frac{1}{a+1} n_{2l} & k \neq l \end{cases} \end{aligned}$$

The payoff can be written as a matrix as follows, where supermarket 1 chooses rows and supermarket 2 chooses columns:

	c_1	c_2
c_1	$\frac{1}{2} n_{11} + \frac{1}{2}(n_{21} + n_{22})$, $\frac{1}{2} n_{11} + \frac{1}{2}(n_{21} + n_{22})$	$n_{11} + \frac{a}{a+1} n_{21} + \frac{1}{a+1} n_{22}$, $n_{12} + \frac{a}{a+1} n_{22} + \frac{1}{a+1} n_{21}$
c_2	$n_{12} + \frac{a}{a+1} n_{22} + \frac{1}{a+1} n_{21}$, $n_{11} + \frac{a}{a+1} n_{21} + \frac{1}{a+1} n_{22}$	$\frac{1}{2} n_{12} + \frac{1}{2}(n_{21} + n_{22})$, $\frac{1}{2} n_{12} + \frac{1}{2}(n_{21} + n_{22})$

Thus, the supermarkets' choice depend on the specific values of the game's parameters, n_{sk} and a. For instance, for a close to 1, $n_{11} = 100$ and $n_{12} = 10$, then both players choose the same constituency, c_1 and they share the market. However, if $n_{11} = 100$ and $n_{12} = 90$, we obtain the Battle of Sexes' payoff matrix, where there are two equilibria, such that in both of them the supermarkets choose different locations.

3 The model

Consider an economy with a finite number of retailers and consumers. Each firm produces or distributes an homogeneous good at zero marginal cost whose price is the same, independently of the firm which provides it. The set of retailers is denoted by R and for sake of simplicity we consider that $R = \{1, 2\}$.

Let C be the set of market places or constituencies, where consumers live and retailers can be located. Let us denote by c_r the constituency where retailer r is located. Let d be a real function on $C \times C$, satisfying that $\forall c, c', c'' \in C$ (i) $d(c, c') = d(c', c)$, (ii) $d(c, c'') \leq d(c, c') + d(c', c'')$ and (iii) $d(c, c) > 0$. Although these properties only guaranty that d is a pseudo-distance, we refer to d as the **distance function** among constituencies.

Let n be the total number of consumers of types Θ. Types summarise all the relevant characteristics of each consumer such as sex, marital status, educational level, economic income, etc. For each $c \in C$, let $\pi_c : \Theta \to [0, 1]$ be the probability distribution of consumers types and by n_c the number of consumers living in constituency c, such that $n = \sum_{c \in C} n_c$.

The two-stage decision problem is structured in the following events. First, each firm $i = 1$ chooses an allocation, simultaneously or following a sequence (the location game). In the analysis we will consider two cases, one in which the location choice is made simultaneously, and other in which we assume that the number assigned to a firm corresponds to its rank in the election. Once both retailers are located, each consumer observes the vector of locations (c_1, c_2), and decides where to buy the good. Given a retailers' location pattern (c_1, c_2), consumers of different types living in the same constituency could buy at different retailers. Let $p_c(c_i|\theta, c_j)$ be the probability of a consumer of type θ living in constituency c to buy to a retailer located in constituency c_i when the other one is located in c_j.

Finally, let Q_c^r denote the amount of sales of retailer $r = 1, 2$ in constituency c and $Q^r = \sum_{c \in C} Q_c^r$. Since each consumer buy a unit of the good, we have that: $Q_c^1 + Q_c^2 = n_c$ and $Q^1 + Q^2 = n$.

In the next subsections we formalise the two stages of the decision problem:

3.1 First stage: location game

In this steps, we have the model has the following elements:

- A set of two players R
- A set of pure strategies for each player, consisting of the set of feasible potential opening constituencies C.
- Payoff functions that assign to a location pattern (c_1, c_2) the market share obtained by the retailers. For each $r = 1, 2$, $r' \neq r$, these functions are given by

$$Q^r(c_1, c_2) = \begin{cases} \frac{1}{2}n & \text{if } \ c_1 = c_2 \\ \sum_{c \in C} n_c \left[\sum_{\theta \in \Theta} \pi_c(\theta) \mu_c(c_r | \theta, c_{r'}) \right] & \text{if } \ c_1 \neq c_2 \end{cases}$$

Note that if $c_1 = c_2$ all consumers choose the same location to buy the good and, since both retailers are identical including prices, consumers choose each establishment with equal probability obtaining each firm a share of 50% of the whole market. However, if firms choose different locations, then a consumer of type θ living in c will buy the good to the firm at r with probability $\mu_c(c_r | \theta, c_{r'})$ depending on the alternative locations to buy.

We can define the normalized market share of retailer r as

$$MS^r(c_1, c_2) = \frac{1}{n} Q^r(c_1, c_2) - \frac{1}{2} \qquad (1)$$

such that $MS^r(c_1, c_2) \in [-\frac{1}{2}, \frac{1}{2}]$ and represents the difference between retailer r's market share and a market share of 50%. We refer to $MS^r(c_1, c_2)$ as retailer r's market share hereafter. Notice that, after normalization, we have that $MS^1(c_1, c_2) + MS^2(c_1, c_2) = 0$ and the location game can be considered as a zero-sum game. Moreover, since retailers can guaranty 50% of the total amount of sales just by choosing the same constituency that their competitor, it is clear that the value of the game, V, is 0.

Von Neumann's Mini-Max theorem guarantees the existence of mixed mini-max strategies for each retailer that provides them with a market share equal to the value of the zero-sum game. Let $\Delta(C)$ be simplex consisting on the distribution space over C, then there exists two probability distributions over constituencies $\sigma_1, \sigma_2 \in \Delta(C)$ such that

$$\begin{aligned} V &= \min_{\sigma_2 \in \Delta(C)} \max_{\sigma_1 \in \Delta(C)} \sum_{c \in C} \sum_{c' \in C} \sigma_1(c)\sigma_2(c') MS^1(c, c') \\ &= \max_{\sigma_1 \in \Delta(C)} \min_{\sigma_2 \in \Delta(C)} \sum_{c \in C} \sum_{c' \in C} \sigma_1(c)\sigma_2(c') MS^1(c, c') = 0 \end{aligned}$$

An important consequence of modelling the location game as a zero-sum game is that the timing in which both retailer inaugurate their premises does not matter. In other words if the leader retailer, say number 1, is the first one to choose location, the apparent advantage for being the first one to decide is illusory: the follower can immediately obtain half of the sales (a normalised market share of 0) just by opening its premises in the constituency selected by the leader. In fact, the follower may choose other location that provides it with an strictly positive normalised market share.

Another important remark is that mini-max theorem guarantees the existence of equilibrium in mixed strategies. Since, in the empirical analysis, it is only possible to observe the constituency actually selected by retailers, we need to refer to equilibrium in pure strategies. We will see, however, that in the considered examples there exists a unique equilibrium in pure strategies.

3.2 Second stage: consumption

We analyze consumer behavior within the Random Utility Maximization paradigm introduced by McFadden (1977). To this end, we assume that each player i of type θ living at constituency c has a utility function on C that summarises her preference of buying a unit of the good at each constituency in C. This utility function has a systematic part, depending on consumer's type and the distance between home and the retailer location, and a random term that summarizes the effects of other variables that could not be measured. Specifically, we assume that the utility of consumer i of type θ living in c to buy in a retailer located in c_r is given by:

$$u_i(\theta, c, c_r) = a(\theta)d(c, c_r)^{-b(\theta)} + \varepsilon_i$$

where $a(\theta)$ and $b(\theta)$ are functions of the type that determines the impact of distance on utility and ε_i is a random variable, with expected value 0 and unknown variance, independent and identically distributed (i.i.d.) among consumers.

Note that higher values of $b(\theta)$ make farther locations less attractive. Most of theoretical models follow the suggestion of Huff (1963) and they fix $b(\theta) = 2$ for all the consumers, independently of their types. However, this is a strong and not very realistic assumption, since perception of distance depends on characteristics of the consumer such as age or owning or not a vehicle.

Since perturbations are i.i.d, for a representative consumer living in c and type θ the utility function is given by

$$u_c(\theta, c') = a(\theta)d(c, c')^{-b(\theta)}$$

Now, the probability of a consumer to buy in a specific retailer depends on the whole location pattern (c_1, c_2). By considering Luce's Utility Axiom (Luce, 1959), the probability to buy in retailer r is given by the ratio of the utility provided by the purchase in retailer r over the sum of utilities of buying in both retailers. Formally, for $r, r' = 1, 2$ and $r \neq r'$, we have that:

$$p_c(c_r|\theta, c_{r'}) = \frac{u_c(\theta, c_r)}{u_c(\theta, c_1) + u_c(\theta, c_2)} = \frac{d(c, c_r)^{-b(\theta)}}{d(c, c_1)^{-b(\theta)} + d(c, c_2)^{-b(\theta)}}$$

That may be written as

$$\begin{aligned} p_c(c_1|\theta, c_2) &= \frac{1}{1 + [d(c, c_2)/d(c, c_1)]^{-b(\theta)}} = \frac{1}{1 + \delta_c^{-b(\theta)}} \\ p_c(c_2|\theta, c_1) &= \frac{[d(c, c_2)/d(c, c_1)]^{-b(\theta)}}{1 + [d(c, c_2)/d(c, c_1)]^{-b(\theta)}} = \frac{\delta_c^{-b(\theta)}}{1 + \delta_c^{-b(\theta)}} \end{aligned}$$

where $\delta_c = \frac{d(c,c_2)}{d(c,c_1)}$ is the relative distance of constituency c to the locations of both retailers. Hence, Luce's Axiom implies, for our utility function, that the probability of a consumer to buy in a specific retailer is a function of the relative distance between retailers. Function $b(\theta)$ summarises the impact of the characteristics of the consumer on her perception of relative distance, and influences the probability to buy in each retailer.

Note that, for this expression of the probability distribution μ_c, the sales of retailer r under location pattern (c_1, c_2) is given by

$$Q^r(c_1, c_2) = \sum_{c \in C} \sum_{\theta \in \Theta} n_c \pi_c(\theta) \frac{d(c, c_r)^{-b(\theta)}}{d(c, c_1)^{-b(\theta)} + d(c, c_2)^{-b(\theta)}} \quad (2)$$

3.2.1 The functional form of $b(\theta)$

Expression (2) shows the total sales of each retailer given a location pattern, a population distribution and a distance function. All the elements in this expression are measured and known, but the specific form of the function relating the attitude towards relative distance of a consumer of type θ, $\delta_c^{-b(\theta)}$ needs to be established. To this end, some assumptions are made on the functional form of $b(\theta)$. First, consider that type θ refers to a one-dimensional characteristic of the consumer. Being sex, age and educational level the available characteristics at the constituency level, since $b(\theta)$ is related to the consumers' mobility, we choose θ as the characteristic potentially more related with attitude towards mobility, i.e. age. Hence, we may think of $b(\theta)$ as a function relating the age of the consumer to her perception of (relative) distance between the location of the two retailers.

We will assume that there are two different attitudes towards distance depending on age: the attitude of young people, less reluctant to walking or driving a longer distance to go shopping, and older people, more reluctant to distance. It is highly probable that most of the young people buy in their working-area rather than on the living-area and may choose a retailer farther away if it suits them. Older people use to buy several times along the week, but do not commute out of their living-area.

Then, to simplify the model we consider the existence of two type of consumers $\Theta = \{\theta^y, \theta^o\}$ where θ^y refers to young consumers and type θ^o to the older ones. Let us denote by $\pi_c^y = \pi_c(\theta^y)$ and $(1 - \pi_c^y) = \pi_c(\theta^o)$. Finally, assume the following functional form for $b(\theta)$:

$$b(\theta) = \begin{cases} -2 + \beta & \text{if } \theta = \theta^y \\ -2 - \beta & \text{if } \theta = \theta^o \end{cases}$$

to reproduce a model close to that of Huff (1963) where the attitude towards mobility varies with the age of the consumer. Under this assumption, the probability of a representative consumer living in c to buy in retailer c_1 is given by:

$$
\begin{aligned}
p_c(c_1|\theta^y, c_2) &= p_c(c_1|\theta^y, \delta_c) = \frac{1}{1+{\delta_c}^{-2+\beta}} \\
p_c(c_1|\theta^o, c_2) &= p_c(c_1|\theta^o, \delta_c) = \frac{1}{1+{\delta_c}^{-2-\beta}}
\end{aligned}
$$

for young and older consumers respectively, where $\delta_c = \frac{d(c,c_2)}{d(c,c_1)}$. Figure 2 shows the shape of the probability function when $\beta = .5$ as a function of the distance ratio δ_c, which corresponds to a S-shaped function.

It can be checked that

$$
\begin{aligned}
\frac{d(p_c(c_1|\theta^y, \delta_c))}{d\delta_c} &= (2-\beta)\frac{{\delta_c}^{-3+\beta}}{(1+{\delta_c}^{-2+\beta})^2} \\
\frac{d(p_c(c_1|\theta^o, \delta_c))}{d\delta_c} &= (2+\beta)\frac{{\delta_c}^{-3-\beta}}{(1+{\delta_c}^{-2-\beta})^2}
\end{aligned}
$$

are positive whenever $\beta < 2$. Moreover $p_c(c_1|\theta^y, \delta_c) = p_c(c_1|\theta^o, \delta_c)$ if and only if $\delta_c = 0$ or $\delta_c = 1$. Finally, note that

$$
\left.\frac{d(p_c(c_1|\theta^y, \delta_c))}{d\delta_c}\right|_{\delta_c=1} = \frac{1}{2} - \frac{1}{4}\beta
$$

and

$$
\left.\frac{d(p_c(c_1|\theta^o, \delta_c))}{d\delta_c}\right|_{\delta_c=1} = \frac{1}{2} + \frac{1}{4}\beta
$$

In conclusion, independently of β,

$$
\begin{aligned}
p_c(c_1|\theta^y, \delta_c) &> p_c(c_1|\theta^o, \delta_c) \text{ if } \delta_c > 1 \text{ and} \\
p_c(c_1|\theta^y, \delta_c) &< p_c(c_1|\theta^o, \delta_c) \text{ if } 0 < \delta_c < 1
\end{aligned}
$$

Hence, when both retailers are at the same distance of a consumer, $\delta_c = 1$, the slope of the probability function is greater for older consumers than for younger ones showing an bigger aversion to the distance for the formers, i.e. for a given increase in the ratio between the distances $d(c, c_2)$ and $d(c, c_1)$, the increase in the probability to buy to the closer retailer in c_1 is bigger for consumer of type θ^o than for consumer of type θ^y.

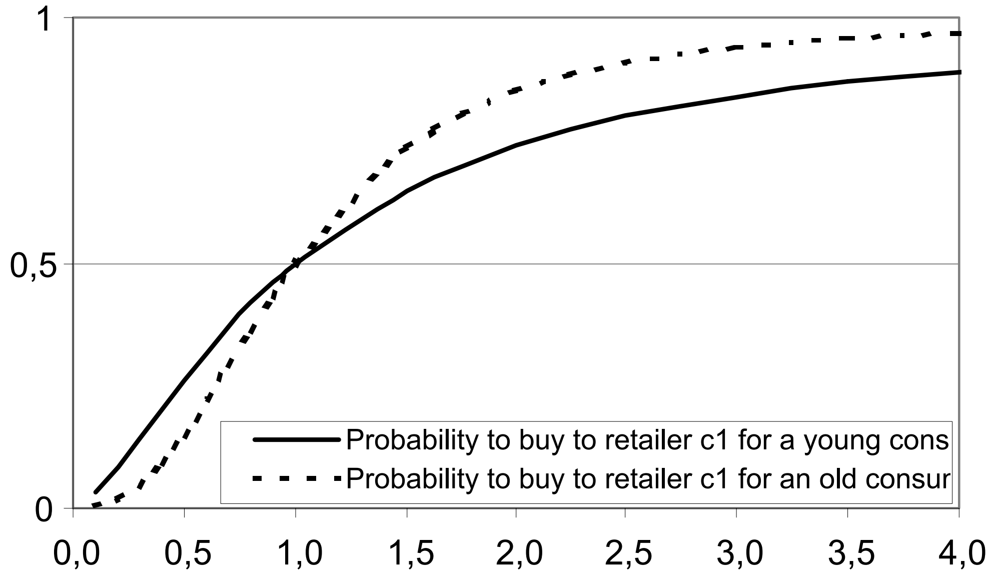

Figure 2: The continuous line shows the probability of choosing retailer c_1 as a function of the distance ratio δ_c for young consumers. The dashed line shows that probability for older consumers. Both lines cross at $\delta_c = 1$, where consumers are indifferent between both retailers.

4 An empirical application to supermarkets location

The goal of this section is the empirical application of the above model to three examples of supermarket location in the Region of Valencia. Specifically, our goal is the obtention of a simple estimate of parameter β and to check whether the actual location of supermarkets gives the value of the location zero-sum game. In this section, we present first the sources of secondary data and the methodology used to build the distance functions. Then, we analyze the estimation of parameter β, calculate the payoff matrix for the location game and compare the actual locations with theoretical equilibria.

We present the case of three middle-sized towns of the Region of Valencia: Requena, Utiel and Segorbe, where in each of them there is one and only one supermarket of both brands, Mercadona and Consum.

4.1 Methodology: secondary sources and field work

For each of the three towns, the following information items were used:

1. A list and map of its constituencies,
2. The specific constituency where each supermarket is actually located,
3. The number of young (between 20 and 54 years) and older (more than 54 years) inhabitants living in these areas,
4. A distance function between constituencies,
5. The market share of Mercadona and Consum in each constituency.

The information referred in points 1, 2 and 3 was obtained from geographical and statistical sources of the National Statistical Institute, INE. The methodology followed to obtain information items 4 and 5 requires further explanation.

4.1.1 Construction of the distance function

Cartographic information at constituency level is not still available in a digital format for all constituencies in Spain. Thus, to construct this function, we proceeded to delimitate the border of all the constituencies and to locate their geographical centroids on a cartographic map of each town. The distance between two different constituencies within the urban center was estimated as the walking time between their centroids. If at least one of the constituencies is located out of the urban center, distances are calculated as driving plus parking times. On the other hand, since we assumed that the distance function was such that $d(c, c) \neq 0$, we calculated the distance from a constituency to itself as the average walking time from the centroid to the vertices of the polygonal formed by its border, when the constituency is in the urban center. When dealing with areas out of the urban center, which consist of a set of disseminated and very small population nuclei, we measured their internal distance as the average driving and parking time from the centroid to the farthest inhabited points. Summarising, we considered:

$$d(c,c') = \begin{cases} \text{walking time between centroids} & \text{if } c \neq c';\ c \text{ and } c' \text{ urban} \\ \text{driving time between centroids} & \text{if } c \neq c';\ c \text{ or } c' \text{ not urban} \\ \text{average walking time from centroid to vertices} & \text{if } c = c';\ c \text{ urban} \\ \text{average driving time from centroid to farthest points} & \text{if } c = c';\ c \text{ not urban} \end{cases}$$

Estimates of the distance matrices are shown in the Appendix.

4.1.2 Estimation of market share

Information on sales of each supermarket is not available at the constituency level. Hence, we estimated a proxy of market share defined in terms of customer shares at a specific time. We assumed that, at any time, the average expenditure of the customers queuing in Mercadona is equal to that of customers queuing in Consum. Then, the ratio of customers in waiting line can be considered as a proxy of the ratio of the sales of both supermarkets.

We measured these ratios for each town every fifteen minutes during an hour (from 12:00 to 13:00) four different days in two weeks (Wednesday and Saturday of the first week and Tuesday and Thursday of the second) in May 2003. We obtained 16 measures of market share in each village. The empirical application uses the average estimate as the actual market share of both supermarkets.

The limitation of secondary demographic data, geographical sources and market information should be overcome by further research.

4.2 Estimation of β

Parameter β summarises the impact of age in the perception of relative distances. Since we are dealing with middle-size towns in the same geographic area, we assumed that this impact must be the same in the three places. To estimate β, we need know supermarkets' location and estimate the empirical market shares in the three towns. These values are given by:

	Requena		Segorbe		Utiel	
	Location	Share	Location	Share	Location	Share
Mercadona	27	0.45	14	0.60	31	0.59
Consum	22	0.55	14	0.40	23	0.41

Let us denote Mercadona as supermarket 1 and Consum as supermarket 2. Our theoretical model derives the total sales of each supermarket $r = 1, 2$, denoted by Q^r, given the distance function and the population distribution, as a function of the aversion to relative distance β. Hence, we can write $Q^r = Q^r(\beta)$. Let $MS^r(\beta) = \frac{1}{n}Q^r(\beta) - \frac{1}{2}$ be the market share of supermarket r. It is clear that $MS^1(\beta) + MS^2(\beta) = 0$, so that the analysis can be limited to one of the supermarkets, say number 1 (Mercadona). From our model, we have that:

$$\begin{aligned} MS^1(\beta) &= -\frac{1}{2} + \frac{1}{n}\sum_{c\in C} n_c[\pi_c^y \frac{d(c,c_1)^{-2+\beta}}{d(c,c_1)^{-2+\beta} + d(c,c_2)^{-2+\beta}} \\ &\quad +(1-\pi_c^y)\frac{d(c,c_1)^{-2-\beta}}{d(c,c_1)^{-2-\beta} + d(c,c_2)^{-2-\beta}}] \end{aligned}$$

or, in terms of relative distance $\delta_c = \frac{d(c,c_2)}{d(c,c_1)}$,

$$MS^1(\beta) = -\frac{1}{2} + \frac{1}{n}\sum_{c\in C} n_c[\pi_c^y \frac{1}{1+\delta_c^{-2+\beta}} + (1-\pi_c^y)\frac{1}{1+\delta_c^{-2-\beta}}]$$

A comparison of this expression with the observed value of the market share, provides an equation for β, for each town, whose supermarkets are located in different constituencies[5]. These equations may be solved by numerical techniques, providing an estimate of the parameter β. Specifically, we have that:

$$\begin{aligned} MS^1(\widehat{\beta}_{requena}) &= -0.05 \Longrightarrow \widehat{\beta}_{requena} = 0.61 \\ MS^1(\widehat{\beta}_{utiel}) &= 0.09 \Longrightarrow \widehat{\beta}_{utiel} = 0.72 \end{aligned}$$

[5] If both supermarkets are located in the same constituency, relative distances are one and the model result is a market share of 0.5 for each supermarket. That is the case of Segorbe in our application.

As expected from the interpretation of the aversion to relative distance, β is positive and its estimates in the two towns are close, supporting our hypothesis of this parameter being constant in the middle-sized towns in the interior of the Region of Valencia. In the following, we estimate β by the average of the estimates obtained in both towns:

$$\widehat{\beta} = \frac{\widehat{\beta}_{requena} + \widehat{\beta}_{utiel}}{2} = 0.67$$

4.3 Payoff matrices and equilibrium analysis

We are interested in finding out pure equilibrium strategies of the location game. Although these strategies do not necessarily exist, we will check that there are one and only one of then for each of the three towns. Since the value of the game is zero, in equilibrium both brands obtain a normalized market share of 0 or, in other words, a 50% of the total amount of consumers' purchases. Hence, equilibrium location strategies consist of selecting a location pattern such that (1) each brand gets half of the total market of consumers and (2) given the location of its competitor, there is no constituency where the other supermarket could get more that 50% of all this market.

In the next subsections we analyze forecastings for the three villages. To make the analysis clear, we only consider as potential opening constituencies those in the urban center. In other words, we assume that a supermarket brand does not choose small disseminate villages for its premises.

4.3.1 Requena

Applying expression (1) and (2) where $\widehat{\beta} = 0.67$, we have that in Requena (normalized) market shares for supermarket 1 are given by the following matrix, where the row refers to the constituency code where supermarket 1 is located and the columns to the constituency where supermarket 2 opens its premises:

Table 1: Payoff Matrix (normalized) for Requena

	11	**12**	**13**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
11	$0,00$	$0,04$	$0,01$	$0,11$	$0,13$	$0,02$	$0,12$	$0,07$	$-0,03$	$0,08$
12	$-0,04$	$0,00$	$-0,03$	$0,06$	$0,09$	$-0,01$	$0,09$	$0,03$	$-0,05$	$0,04$
13	$-0,01$	$0,03$	$0,00$	$0,07$	$0,10$	$0,02$	$0,12$	$0,05$	$-0,03$	$0,07$
21	$-0,11$	$-0,06$	$-0,07$	$0,00$	$0,03$	$-0,07$	$0,03$	$-0,03$	$-0,10$	$-0,01$
22	$-0,13$	$-0,09$	$-0,10$	$-0,03$	$0,00$	$-0,10$	$0,00$	$-0,06$	$-0,15$	$-0,05$
23	$-0,02$	$0,01$	$-0,02$	$0,07$	$0,10$	$0,00$	$0,09$	$0,05$	$-0,06$	$0,05$
24	$-0,12$	$-0,09$	$-0,12$	$-0,03$	$0,00$	$-0,09$	$0,00$	$-0,05$	$-0,15$	$-0,06$
25	$-0,07$	$-0,03$	$-0,05$	$0,03$	$0,06$	$-0,05$	$0,05$	$0,00$	$-0,09$	$0,01$
26	$0,03$	$0,05$	$0,03$	$0,10$	$0,15$	$0,06$	$0,15$	$0,09$	$0,00$	$0,10$
27	$-0,08$	$-0,04$	$-0,07$	$0,01$	$0,05$	$-0,05$	$0,06$	$-0,01$	$-0,10$	$0,00$

Notice that, since we are dealing with a zero-sum game, supermarket 1 cannot obtain an advantage for the fact of being the first to choose location. Hence, this payoff matrix can be used to analyze the problems of both simultaneous and sequential opening.

From matrix in table 1, it is very easy to check that there is only one equilibrium in pure strategies, that consisting of both supermarkets choosing constituency 26. To this end, notice that if supermarket 1 chooses the row corresponding to constituency 26, supermarket 2 will choose the column that guaranties the highest market share for it. Since this matrix shows payoffs for supermarket 1 and the sum of both shares is zero, the best row for supermarket 2 is that showing the lowest share for supermarket 1: constituency 26, where supermarket 1 gets null payoff. Of course, in equilibrium, both firms obtains a market share of 0 (the value of the game).

Note that actual locations are given by leader supermarket (Consum) in constituency 22 and the follower (Mercadona) in constituency 27. Even if this location pattern differs from the equilibrium outcome, the payoff matrix helps us to understand the process. Let us assume that Consum chooses its location in terms of considerations out of the model. To create a market in a town of the characteristics of Requena, Consum has located its premises in the very center of the town, close to the central market. Once this location is public, the follower chooses the constituency that maximizes it market share given its competitor location. We can see in the payoff matrix that this is constituency 26. Although it does not fit exactly with the actual location of Mercadona, we can see in the map that it is placed quite close to this area.

4.3.2 Utiel

Payoffs for supermarket 1 in the location game in Utiel are given by matrix in table 2, whose rows correspond to selection of supermarket 1 and columns to those of supermarket 2:

Table 2: Payoff Matrix (normalized) for Utiel

	11	**12**	**13**	**21**	**22**	**23**	**31**	**32**	**33**
11	$0,00$	$-0,05$	$0,00$	$-0,17$	$-0,10$	$-0,17$	$-0,06$	$-0,18$	$-0,15$
12	$0,05$	$0,00$	$0,04$	$-0,12$	$-0,06$	$-0,12$	$-0,01$	$-0,14$	$-0,10$
13	$0,00$	$-0,04$	$0,00$	$0,06$	$-0,11$	$-0,15$	$-0,05$	$-0,16$	$-0,14$
21	$0,17$	$0,12$	$0,18$	$0,00$	$0,09$	$0,02$	$0,10$	$-0,01$	$0,02$
22	$0,10$	$0,06$	$0,11$	$-0,09$	$0,00$	$-0,06$	$0,04$	$-0,07$	$-0,05$
23	$0,17$	$0,12$	$0,15$	$-0,02$	$0,06$	$0,00$	$0,09$	$-0,03$	$0,00$
31	$0,06$	$0,01$	$0,05$	$-0,10$	$-0,04$	$-0,09$	$0,00$	$-0,15$	$-0,11$
32	$0,18$	$0,14$	$0,16$	$0,01$	$0,07$	$0,03$	$0,15$	$0,00$	$0,03$
33	$0,15$	$0,10$	$0,14$	$-0,02$	$0,05$	$0,00$	$0,11$	$-0,03$	$0,00$

As in Requena, there exists just one equilibrium in pure strategies, which provides both brands with half of the total market. Equilibrium in Utiel is given by both supermarkets choosing constituency 32 to open their premises.

This forecasting does not fit with actual location pattern in Utiel, where the leader supermarket chose constituency 31 and the follower, Consum, constituency 23. However, we can see that location chosen by the leader is very close geographically to that forecasted in the model, and that selecting 23 is not the best but a good response of Consum, given our payoffs matrix, since it provides Consum with a market share higher than that of the leader Mercadona.

4.3.3 Segorbe

Table 3 shows payoffs matrix for the location game in Segorbe. As in the previous two cases, rows refer to selection of supermarket 1, whose market shares are presented in the table, and columns to the location selection of supermarket 2:

Table 3: Payoff Matrix (normalized) for Segorbe

	11	**12**	**13**	**14**	**15**	**21**	**22**	**31**	**32**
11	$0,00$	$-0,05$	$0,01$	$-0,07$	$-0,08$	$-0,06$	$-0,02$	$-0,03$	$-0,02$
12	$0,05$	$0,00$	$0,05$	$-0,02$	$-0,04$	$-0,02$	$0,02$	$0,02$	$0,03$
13	$-0,01$	$-0,05$	$0,00$	$-0,08$	$-0,09$	$-0,07$	$-0,03$	$-0,03$	$-0,03$
14	$0,07$	$0,02$	$0,08$	$0,00$	$-0,02$	$0,00$	$0,04$	$0,04$	$0,04$
15	$0,08$	$0,04$	$0,09$	$0,02$	$0,00$	$0,01$	$0,05$	$0,04$	$0,06$
21	$0,06$	$0,02$	$0,07$	$0,00$	$-0,01$	$0,00$	$0,04$	$0,04$	$0,05$
22	$0,02$	$-0,02$	$0,03$	$-0,04$	$-0,05$	$-0,04$	$0,00$	$0,04$	$0,00$
31	$0,03$	$-0,02$	$0,03$	$-0,04$	$-0,04$	$-0,04$	$0,00$	$0,00$	$0,01$
32	$0,02$	-0.03	$0,03$	$-0,04$	$-0,06$	$-0,05$	$0,00$	$0,00$	$0,00$

As in the previous cases, there exists only an equilibrium in pure strategies in which both brands select constituency 15 for their premises. Actually, in Segorbe both supermarkets are located in the same constituency, number 14, very closed of the location forecasted for the model. Note that Segorbe is the only of the three cases where selection of constituencies can be considered as simultaneously, since the opening of Mercadona (the first one) and Consum differs in less than a year and brands should have made their decision while ignoring that their competitors were to open premises in Segorbe and, of course, the location chosen for it.

4.4 Conclusions from empirical analysis

We can summarise the conclusions of the empirical analysis of equilibrium in the following points:

- There exist one and only one equilibrium in pure strategies for each of the three towns,
- This equilibrium is always 'pooling': both brands choose the same constituency for their premises,

- In equilibrium, both supermarkets get the value of the game, consisting of half of the total market,
- The follower brand can always obtain at least a half of the total market,
- Payoff matrix partially explains location decisions of the follower brand,
- Forecasted location market is the same, no matter if location are chosen simultaneously or sequentially.

As we can see by comparing these forecasting with the empirical facts, the model seems to work quite good for Segorbe, where decision was made simultaneously. In Requena and Utiel, where decisions were clearly sequential, the model fails to forecast the selection of the leader but it explains, up to some geographical imprecisions, the response of the follower. On the other hand, we observe that leader brand always keeps a bit more than a half of the total market, meanwhile the model forecasts exactly half of the market (in equilibrium) or even a bit less out of equilibrium.

In our opinion, these differences between theoretical equilibrium and empirical actual locations in sequential decisions can be a consequence of one of the assumption of the location game: to be a zero-sum game. We assume that the total purchases of the consumers does not depend on the specific location of supermarkets. However, this could not be the case ten years ago, when leader supermarkets (Consum in Requena and Mercadona in Utiel) opened their premises. Before these opening, all the market where served by a large number of small retailers distributed among all the municipal term. Hence, leader supermarkets choose specific and more appealing locations, such as constituency 22 in Requena, very close to the central market and in the middle of the commercial area of the town that is visited for a large proportion of consumers, independently of the presence of the supermarket. Hence, supermarkets may choose some specific location patterns, looking for and increment of the total sales that consumers buy at supermarkets, instead of at small retailers. The relaxing of the assumption of zero-sum in the location game is one of our current research goals. This relaxation, and the inclusion of some consideration on brand image, seem to be very powerful to improve the model.

5 Concluding remarks

In this paper we present a theoretic two-stage model for retailers location (stage 1) and consumers purchase decision (stage 2). Retailers decision problem is formalized in terms of a zero-sum game, whose payoffs refers to retailers' market share. Consumers decision problem is formalized in terms of a discrete choice model, based on random utilities. These two alternative techniques have been selected to discriminate between the very strategic behavior of a small number of retailers and the less strategic behavior of a large number of consumers. The theoretical model provides forecasting on the equilibrium market share to be obtained (the value of the remaining zero-sum game) and the locations to be chosen by retailers, in terms of the geographic distribution of the underlaying location space (constituencies of the town), population distribution and characteristics (types) of the consumers.

The theoretic model is applied to analyze empirically supermarket distributions on three middle-sized towns of the interior of the region of Valencia: Requena, Segorbe and Utiel. Departing from actual market shares in these three towns, a function relating types (young and older consumers) and sales is estimated. This estimation allows us to calculate the payoff matrices for the location games corresponding to the three towns and to find out that there is just one equilibrium in pure strategies for each town. A comparison of the equilibrium and the actual situation of the supermarkets shows that our model explain quite well decision location but it must be improved to obtain more precise forecastings of actual market shares.

We consider that this improvement can be obtained by relaxing one of the key assumptions in the model: the fact of the location game being a zero-sum one. It can be assumed that the total purchases depend on some characteristics of the constituencies where retailers are located and/or the aggregate distance they must walk or drive to go from home to supermarkets. Another improvement point to be analyzed is the assumption that, since supermarkets have equal prices, consumers are indifferent between buying in each of them. Since leader supermarket is the first one to open its premises, it may create consumption habits or brand considerations that have not been taken into account in the model, and could be the responsible of the leader having higher market shares. On the other hand, our model should be enriched to collect some facts of consumer behavior such as that they do not overpass easily a location with a supermarket to go to another one, or the tendency to go from outer neigborhoods to the urban center more than from the center to the outer zone (asymmetric distance function). The design of a richer model, including all these considerations, and to test it empirically in a wider set of towns is our goal in our future research agenda.

6 Bibliography

Brenner, S. (2001). 'Determinants of product differentiation: A Survey'. Humboldt University Research paper, Mimeo.

Chasco, M.C. (1997). 'Modelos de Determinación de Áreas de Mercado del Comercio al Por Menor'. Instituto Klein Research paper, Mimeo.

Fotheringham, A.S. and M.E. O'Kelly (1989). *Spatial Interaction Models: Formulations and Applications.* Kluwer Academic Publishers. Studies in Operational Science.

Hotelling, H. (1929). 'Stability in Competition', *Economic Journal,* **39**, 41-47.

Huff, D.L. (1963). 'A Probabilistic Analysis of Consumer Behavior'. William S. Decker (Ed.), *Emerging Concepts in Marketing.* 443-461, American Marketing Association, Chicago.

Louvière, J. and G. Woodworth, G (1983). 'Design and Analysis of Simulated Consumer Choice of Allocation Experiments: An Approach Based on Aggregate Data'. *Journal of Marketing Research*, **20**, 350-367.

Luce, R. (1959). *Individual choice behavior: a theoretical analysis*, J. Wiley and Sons, New York

McFadden, D.(1977) 'Econometric Models of Probabilistic Choice'. C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, Mass.: MIT Press.

7 Appendix: empirical data

In Spanish statistical system, constituencies are denoted by correlative one digit numbers within each district. On the other hand, district are named in the say way and we can refer to a constituency as a two digits number, the first one referring to the district and the second to the constituency number in this district.

Requena

The municipal term of Requena is divided in fifteen constituencies. Constituencies 11, 12, 13, 21, 22, 23, 24, 25, 26 and 27 correspond to the urban center and 31, 32, 33, 41 and 42 to clusters of very small villages around the main urban center. Population and pseudodistances of these constituencies are given by:

Constituency	Population	% young inhabitants
11	1.713	54
12	1.617	59
13	737	54
21	593	50
22	1.011	61
23	1.274	69
24	1.022	64
25	1.263	64
26	1.657	68
27	1.847	74
31	1.748	57
32	682	44
33	642	46
41	551	47
42	518	40
Total	17.075	61

	11	12	13	21	22	23	24
11	7.6						
12	7.6	3.8					
13	11.4	4.8	3.8				
21	4.8	9.5	12.4	2.9			
22	8.6	9.5	11.4	3.8	3.8		
23	14.3	15.2	17.1	8.6	6.7	4.8	
24	9.5	7.6	8.6	6.7	2.9	8.6	4.8
25	9.5	12.4	15.2	3.8	3.8	4.8	6.7
26	19.0	17.1	16.2	13.3	9.5	6.7	8.6
27	11.4	7.6	5.7	9.5	5.7	11.4	2.9
31	21.0	21.0	18.1	21.0	21.0	21.0	18.1
32	15.2	15.2	18.1	15.2	15.2	15.2	18.1
33	21.0	21.0	23.8	21.0	21.0	21.0	23.8
41	27.6	27.6	30.5	27.6	27.6	27.6	30.5
42	25.7	25.7	28.6	25.7	25.7	25.7	28.6

	25	26	27	31	32	33	41	42
25	3.8							
26	11.4	5.7						
27	10.5	9.5	4.8					
31	21.0	18.1	18.1	9.5				
32	15.2	18.1	18.1	24.8	8.6			
33	21.0	23.8	23.8	30.5	13.3	8.6		
41	27.6	30.5	30.5	37.1	20.0	6.7	4.8	
42	25.7	28.6	28.6	35.2	18.1	16.2	21.9	9.5

Utiel

The municipal term of Utiel is divided in eleven constituencies, where constituencies 11, 12, 13, 21, 22, 23, 31, 32 and 33 correspond to the urban center and 41 and 42 to small villages near Utiel. Population and pseudodistances of these constituencies are given by:

Constituency	Population	% young inhabitants
11	713	53
12	636	64
13	555	47
21	912	61
22	1.233	60
23	1.601	64
31	1.152	63
32	1.202	64
33	1.238	62
41	613	45
42	768	48
Total	10.623	59

	11	12	13	21	22	23	31	32	32	41	42
11	4.5										
12	2.7	5.5									
13	5.5	8.2	3.6								
21	13.6	18.2	10.0	12.7							
22	8.2	10.9	5.5	7.2	5.5						
23	10.9	10.0	10.9	13.6	16.4	8.2					
31	7.3	8.2	8.2	18.2	13.6	16.4	5.5				
32	12.7	12.7	13.6	14.5	20.0	20.0	7.3	11.8			
33	14.5	15.5	11.8	15.5	16.4	21.8	8.2	13.6	7.3		
41	18.2	15.5	18.2	18.2	18.2	15.5	18.2	18.2	18.2	5.5	
42	9.1	15.5	18.2	18.2	18.2	15.5	18.2	18.2	18.2	14.5	11.8

Segorbe

Segorbe is divided in nine constituencies, all of them corresponding to the village, 11, 12, 13, 14, 15, 21, 22, 31 and 32. Population and pseudodistances of these constituencies are given by:

Constituency	Population	% young inhabitants
11	716	59
12	674	51
13	884	55
14	1.674	67
15	608	77
21	602	49
22	679	59
31	616	43
32	668	48
Total	7.121	58

	11	12	13	14	15	21	22	31	32
11	6.4								
12	3.6	4.5							
13	3.6	7.3	6.8						
14	10.0	13.6	6.8	8.2					
15	10.5	14.1	7.3	5.9	6.4				
21	12.7	13.6	12.7	13.2	17.7	3.6			
22	10.0	11.8	9.1	8.6	13.6	4.5	5.9		
31	9.5	9.5	10.5	12.7	16.8	4.1	5.5	4.1	
32	6.8	5.5	8.6	13.6	15.9	8.2	7.7	4.1	3.6