The Follower Competitive Location Problem with Comparison-Shopping

Vladimir Marianov[1] · H. A. Eiselt[2] · Armin Lüer-Villagra[3]

Abstract

In competitive settings, firms locate their stores to take advantage of consumers' behavior to maximize their market share. A common behavior is comparison-shopping: in this behavioral pattern, consumers visit multiple stores that sell non-identical products, which are mutual substitutes, before making their purchase decision. This behavior has never been included in location-prescribing models for competitive firms. Given existing branches of one firm, we address the location problem of a follower firm that locates its own branches. We present insights on the instance used by ReVelle in his *maximum capture* formulation, provide computational experience with one thousand 100-node instances, and consider a realistic case using a 353-node network of Santiago, Chile. The results are compared in terms of the demand captured by each firm and the locational patterns that result from different consumer behaviors.

Keywords Competitive location · Comparison-shopping · Follower problem

1 Introduction

Facility location researchers and practitioners deal with finding the best possible locations for all kinds of infrastructure. The book by Laporte et al. (2015) offers a

✉   Vladimir Marianov
      marianov@ing.puc.cl

      H. A. Eiselt
      haeiselt@unb.ca

      Armin Lüer-Villagra
      armin.luer@unab.cl

1   Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna, 4860 Santiago, Chile

2   Faculty of Business Administration, University of New Brunswick, Fredericton, NB, Canada

3   Engineering Science Department, Universidad Andres Bello, Santiago, Chile

comprehensive view of the field while the contributions in Eiselt and Marianov (2015) discuss selected examples of applications. An important area in the field of location theory deals with competitive scenarios. In his seminal paper, Hotelling (1929) studied a market in the shape of a line segment (a so-called "linear market") with uniformly distributed demand, on which two competing facilities locate and sell the same product. Both firms locate a single facility each, and their decision variables are the location of their facility as well as the mill prices they charge. Customers are assumed to patronize the facility at which they will pay the lowest full price, i.e., the mill price plus the transportation costs, which are assumed to be linear in the distance. The main thrust of Hotelling's work dealt with the question whether or not there is an equilibrium in the situation he described. A (Nash) equilibrium is defined as a pair of decisions of the competitors, which is stable, i.e., neither competitor can benefit by unilaterally changing his decision. Hotelling's conclusion was that such equilibrium exists with both firm locating their facilities at the center of the market, a result that was dubbed *the principle of minimum differentiation*.

In the early days following Hotelling's paper, a number of his followers considered the principle of minimum differentiation the ultimate explanation for the frequently observed clustering or agglomeration of competitive facilities in practice. However, the decades following Hotelling's paper showed many limitations of the analysis. First, the result was rather fickle similar to the equilibrium a ball is in when located on a plain surface: even a minor change of the slope of the surface will destroy the equilibrium (at least in the absence of friction). Secondly, it turned out that small changes in the assumptions could result in dispersed locations (see, e.g., Chamberlin 1933; Lerner and Singer 1937; Eaton and Lipsey 1975; Okabe and Suzuki 1987; Brown 1989). Thirdly, D'Aspremont et al. (1979) in their paper demonstrated that Hotelling's conclusion with his own assumptions did not hold and that the original problem did not possess an equilibrium.

Another strand of analysis dates back to the work of the economist von Stackelberg (1943). The competitive situations in his work comprise two (classes of) competitors: the one(s) to act first are the *leaders*, while those that act after the leader has made a decision and this decision has become public knowledge, are the *followers*. Once decisions have been made, they are irreversible. Note that the scenario is asymmetric: while the leader has to consider the potential decisions of the follower, the follower does not need such foresight; he only takes the situation as given and optimizes his own objective. It is noteworthy that the follower's problem is a conditional location problem ("optimize your own objective, given that the leader has already located at known sites"), while the leader must include the follower's reaction in each of his decisions, resulting in a bi-level optimization problem (see, e.g., Aras and Küçükaydın 2017). The first to apply such a leader – follower model to competitive location problems (albeit with fixed and equal prices) are Prescott and Visscher (1977). While these authors still worked on Hotelling's linear market, Drezner (1982) considered competitive location problems in the plane, while Hakimi (1983) discussed similar problems on networks. Hakimi also coined the expression *centroid* for the leader's problem and *medianoid* for the follower's problem. Later, ReVelle (1986) solved the follower problem (dubbed the *maximum capture* or MAXCAP problem) on a network using an integer programming formulation. These contributions started a large body of research dedicated to the subject of competitive location by operations researchers, most of it solving one of

the von Stackelberg problems, sometimes considering location and pricing, as in Kress and Pesch (2016). Berglund and Kwon (2014) present a different asymmetric approach, in which a von Stackelberg firm competes with Cournot-Nash firms. For reviews, see, e.g., Eiselt et al. (2015), Kress and Pesch (2012) for problems on networks; and Drezner (2014) for problems in the plane. Eiselt (2011) as well as Marianov and Eiselt (2016) analyze competitive location and agglomeration results from the point of view of location researchers.

One of the key features in any competitive location problem concerns customer behavior. While Hotelling assumed that customers would purchase the good from the source with the lowest full price, most of his successors, who did not include price competition in their models, reduced this assumption to its simplified version, in which customers purchase from the source closest to them. A number of authors have made different assumptions, which essentially fall into two categories. First, there are those authors who have dropped the "individual trip" assumption made by almost all researchers in location analysis. In other words, consumers are not necessarily assumed to make special trips for each product they attempt to purchase. This bundling is typically referred to as multipurpose shopping or, as the case may be, multi-stop shopping. Dellaert et al. (1998) provide some insight into multipurpose shopping behavior from a conceptual point of view, while Hodgson (1990) introduced the concept of flow capturing (or flow interception). In it, customers no longer choose the facility closest to their respective home locations, but a facility closest to a trip they are on anyway, e.g., the daily trip to work. In other words, the relevant customer - facility distances are no longer point-to-point distances, but path-to-point distances. This concept applies when customers handle drop-offs (as in the case of childcare facilities) or pickups (such as gas fill-ups) along the way between home and work. Lately, there has been a revival of this idea when locating alternative fuel stations, as in Miralinaghi et al. (2017). This idea can be seen as a limited version of location-routing problems, see, e.g., Nagy and Salhi (2007). Marianov et al. (2018) investigate the effects of multipurpose shopping on store location.

Another aspect of customer behavior concerns information gathering. In the case of retail location models, this would include internet searches, flyers, ads in media, and information collection by visiting the stores. Most authors in this subfield consider in price search. Contributions such as those by Guo and Lai (2014) include not only internet search, but also internet purchases in their model. The present paper belongs into this category, as we deal with customers comparing products in stores before making any purchase, i.e., comparison-shopping. More specifically, we assume a customer behavior that includes two levels of comparisons: in the first pre-trip planning stage, customers compare observable primary features of the products in which they are interested. This will include price as well as specifications that are quantifiable and typically published, such as the size of an item, its weight, its primary product features, etc. Given the results of the research, the consumer will plan a trip. Then, in the second during-trip stage, customers will examine the secondary product features (such as color, flavor, and specific fit in the case of clothing) in detail and make their decision to purchase or not to purchase accordingly. This behavior will be described in detail in the second section of this paper.

The economic literature has addressed multipurpose and comparison-shopping, as in Eaton and Lipsey (1975, 1979, 1982), McLafferty and Ghosh (1987), Ghosh and

McLafferty (1984), Arentze et al. (2005), Mulligan (1987), O'Kelly (1981, 1983), Thill (1982), and Wolinsky (1983). However, to the best of our knowledge comparison-shopping has never been dealt with in prescriptive facility location models see, e.g., Santos-Peñate et al. (2019) and Pelegrín et al. (2018).

Our contribution to the literature consists in including, for the first time, comparison-shopping in a competitive facility location problem. We propose a model that solves the follower problem for a market-share-maximizing firm locating one or more stores in the presence of existing competitor's stores. The follower problem has been addressed in the literature, and it has an importance by itself: it requires being solved by a new entrant in a market in which one or more incumbents are already present. It is also a sub-problem of the leader problem. Since comparison-shopping is a consumer behavior that has never been considered before in this context, it seems reasonable to solve this simpler problem first, to observe the effects of comparison-shopping without interference by other issues.

We use a duopoly for the sake of simplicity and clean analysis. If more chains are involved, there could be effects related to e.g. the market size and border effects, which become undistinguishable from the effects of comparison-shopping alone.

Our analysis compares the locations and demand captures that result from a model with comparison-shopping to those with single- purpose shopping trips. Using the simplest possible setting, we show that comparison-shopping results in a larger market, more frequent co-location, and stronger agglomeration of competing stores.

There are similarities between the multipurpose shopping problem (Marianov et al. 2018) and the comparison-shopping problem investigated here. The first similarity is that both are in essence bi-level problems that we formulate at once as one-level optimization problems. Secondly, in both cases the stores of different firms tend to be located close to each other. However, there are also significant differences, which are shown in Table 1.

In summary, the multipurpose and comparison shopping models have common properties, but there are significant differences in the concepts behind them. The main

Table 1 Differences between Marianov et al. (2018) and the present paper

Marianov et al. (2018)	Present paper
Cooperative model	Competitive model
Products are *not* mutual substitutes	Products are mutual substitutes
Customers can purchase both products on one trip	Customers purchase at most one product on their trip
Monopolistic markets	Shared markets
One decision stage: visit one store or both	Two decision stages: visit one store or two, then, if two stores were visited, what product to purchase.
No customer choice between products of the same type	Customers have the choice between products of the same type
Certainty a priori about the utilities	Uncertainty about utilities: utility depends on the product choice and, if an on-site inspection of the products is not satisfactory, customers may not make any purchase at all, thus resulting in a negative utility

contributions of both papers are different. In both cases, though, the inclusion of more refined assumptions about customer behavior result in better representations of reality.

The remainder of the paper is organized as follows. In Section 2, we describe the problem. Section 3 develops the model. Section 4 contains the computational experience, and Section 5 presents conclusions and future extensions.

2 The Problem

A firm or chain L (the "leader") has several stores located in a market, at known locations, all of them selling the same product. As long as no confusion can arise, we will refer to the chain and its product as L. A second chain F (the "follower") wants to enter the market by locating one or more stores and selling product F. Products L and F are heterogeneous mutual substitutes, i.e., they differ only in secondary features. If a consumer wants to buy a unit of product, he will purchase L or F, but not both. This paper solves the follower problem, i.e., that of the newcomer firm F wanting to maximize its market, given that the branches of firm L have already been located.

The products are non-essential, meaning that consumers may choose not to make any purchase. In addition, consumers are not willing to spend more than a maximum amount of resources (the reservation price) to acquire a unit of the product. Given that both firms apply mill pricing, the reservation price includes the price consumers pay for the product at the branch of the firm they patronize, and the generalized cost for the round trip, which includes the cost of travel, parking, the customer's value of time, and other features. Rather than single consumers, we assume consumer clusters with similar characteristics.

In the first stage of the customers' decision-making process, customers decide irrevocably to either forego the purchase, make a trip to a single store (single-stop trip – *SST*), or engage in comparison-shopping by visiting two stores in the two possible directions (two-stop trip – *TST*). All of this is decided based on perceived expected utility comparisons. Suppose now that a customer has decided to patronize a single store. After visiting the store and obtaining details of the product in question, he will decide, with some probability α, to make the purchase. Alternatively, he may decide not to make the purchase and return home. We assume that customers do not change their mind along the way in the sense that they, after deciding on a single stop trip, obtain the new information at the store and now decide to engage in comparison-shopping. Wolinsky (1983) assumes the same type of shopping behavior, and Drezner et al. (1996) define a choice rule as *consistent*, if a decision to patronize a store does not change along the way.

Suppose now that by way of the utility function, a customer decides to engage in comparison-shopping. This means that the customer will visit the first store—in the order decided upon in the utility function—followed by a visit to the second store. At this point, the customer has full information about the products. The customer now makes the decision to purchase the product with a given probability $\beta \geq \alpha$, in one of the visited stores. In the second stage, the consumer decides what product to purchase. Note that $\beta \geq \alpha$ because the probability of finding a suitable product (and buying it) in visiting branches of both stores is higher than finding an acceptable product in a single store (Huff 1963; Stahl 1982).

In order to describe the second stage of comparison, it is necessary to analyze product features and the way customers compare them. There are two main classes of features: those that are observable from anywhere and those that are not. Clearly, price is a feature that can easily be observed from anywhere, meaning that customers can sit at home and compare products as far as their respective prices are concerned. Some of the product features are more difficult to ascertain and compare: not all flyers, newspapers, or websites that describe the different products list the same features, and intangibles, such as the "feel" for a product, the fit of a piece of clothing, or the ease to use a product, can only be established at a store. In addition, we consider that consumers have taste uncertainty (their taste can change from one purchase to the next), and they might have different preferences regarding these features or completely dislike some products. These are the reasons why they compare.

The second stage of comparison involves features, not maximum utility. If the maximum expected utility rule were used, there would be no reason to visit two stores instead of one, as a single trip would always have a higher utility than a trip visiting two stores. There would be no room for comparison. Furthermore, at the time of the first decision stage, there is uncertainty about what product will be chosen by the consumer. Thus, we require including the uncertainty and lack of information in the expression of the utility, by assuming that, on starting the trip, the consumers do not know what product they will finally choose. This is a fundamental difference with the contribution by Marianov et al. (2018). In the second stage, the consumer does not know what product he will purchase, because they have different desirable or secondary features, about which he does not have information. He needs to estimate the utility by assigning a percentage of patronage, say 50%, to each competitor. We use this figure, corrected by the cost of each alternative (purchasing at the second visited store or coming back to the first visited store and making the purchase there), as we will see in the next Section. This is the expected utility of the two-store trip.

It is important to mention that performing this estimation using other rules, e.g., discrete choice and gravity models (see, e.g., Hodgson 1978; Drezner and Drezner 1996; Fernández et al. 2007; Marianov et al. 2008) is trivial. The results should not change significantly.

Consumers are located at discrete points of a plane or at nodes of a network representing the region. Each point concentrates a known amount of homogeneous demand although, if there are different demand segments at each point, the point can be replaced by as many copies of itself as demand segments, each copy housing homogeneous demand. Throughout this contribution, we assume that all customers at point i have the same reservation price which indicates the value they assign to, i.e., the maximal price they are prepared to pay for a single unit of the good in question. We use reservation prices associated with the customer location, based on the general income level of the area, as reservation prices indicate the ability and willingness to pay for a product. Also, we use all customer locations as candidate locations for stores of both chains, which can be co-located. Figure 1 shows the problem setting and the representation we use in this paper of the customers and store locations.

As an example of the customers' decision-making process, consider a customer, who is interested in a pair of dress shoes. The shoes would nicely complement a given
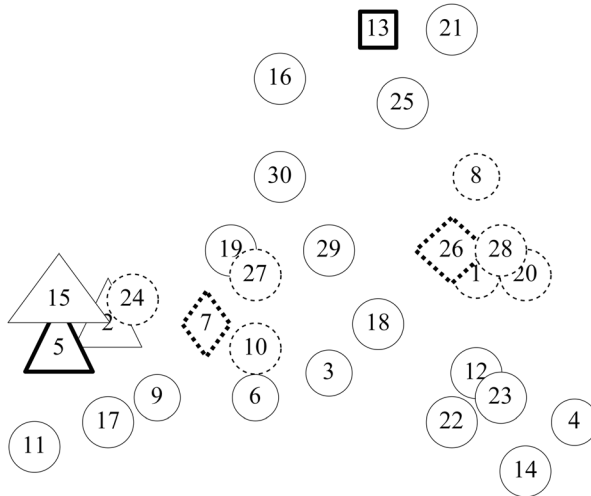
Fig. 1 The problem setting with customers, leader and follower locations. A bold square denotes a leader location, a thin square, customers captured by the leader in single trips. A bold triangle denotes a follower location, a thin triangle, customers captured by the follower in a single trip. Diamonds denote co-location of leader and follower. Circles are customer locations. A dashed circle border denotes they engage in two-stop trips and the demand is shared between both chains

outfit, but their purchase is not necessary to complete it, making them a non-essential good. Based on his location, the customer has a specific reservation price, is aware of the general price level of the two chains and plans accordingly.

3 The Follower Location Model with Comparison-Shopping

For ease of reading, we first define our notation.

Sets

N The set of nodes

S $\{p, q\}$, where p and q take one of the values L or F, i.e., leader store location or product, or follower store location or product.

$I \subseteq N$ The set of demands, indexed by i. Without loss of generality, we assume $I = N$

$J \subseteq N$ The set of candidate locations of the follower's F stores, indexed by j. Without loss of generality, we assume $J = N$

$K \subseteq N$ The set of known locations of the leader's L stores, indexed by k.

Parameters

r_i The reservation price of the consumer $i \in I$.

π^p The price of product p.

g_{ij} The generalized cost of a trip from i to j, with $i, j \in N$

α Probability of finding a suitable product and purchasing it in an *SST*.

β Probability of finding a suitable product and purchasing it by visiting two stores.

ε The largest utility difference that is considered as a computational zero.

u_{ij}^{p} Utility for consumer i of making an *SST* to a store p located at site j.

u_{ijk}^{pq} Utility for consumer i of making an *TST* to a store p located at site j, and then to a competing store q located at site k. We assume that $p \neq q$.

p_{ijk}^{pq} The probability of making a purchase of product q at the second visited store k and going back home. $\left(1 - p_{ijk}^{pq}\right)$ is the probability of returning to the first visited store j and purchasing product p there.

a_{i} The number of customers (a proxy for demand) at node i.

n_{F} The number of stores located by the follower, F.

Additional sets

$\overline{K}_{i}$

$\left\{ k_{t} \in K \,|\, u_{ik_{t}}^{L} \geq u_{ik}^{L}, \forall k \in K \right\}$ set of locations of leader's stores that provide the highest utility to customers at i among all the leader's stores. Let k_{t} be a representative location of this set.

N_{i}^{F}

$\left\{ j \in J \,|\, u_{ij}^{F} \geq u_{ik_{t}}^{L} + \varepsilon \right\}$, set of all candidate locations j for an F store, such that the utility for customers at i for purchasing at j in an *SST*, if a store were located there, is higher than purchasing product L at any $k_{t} \in \overline{K}_{i}$, also in an *SST* trip.

$\overline{N}_{i}^{F}$

$\left\{ j \in J \,|\, u_{ij}^{F} \leq u_{ik_{t}}^{L} - \varepsilon \right\}$, set of locations j whose utility for consumers at i is strictly lower that the utility of purchasing at $k_{t} \in \overline{K}_{i}$, both in SST trips.

$N_{i}^{L=F}$

$\left\{ j \in J \,|\, u_{ij}^{F} \in \left[u_{ik_{t}}^{L} - \varepsilon, u_{ik_{t}}^{L} + \varepsilon \right] \right\}$, set of all candidate locations j for an F store, such that customer i, purchasing at j in an *SST*, has the same utility as purchasing at $k_{t} \in \overline{K}_{i}$.

$$N_{i}^{F} \cup \overline{N}_{i}^{F} \cup N_{i}^{L=F} = J$$

Decision variables

x_{j} One if an F store is located at j, and zero otherwise. The only location variable.

y_{i} One if for the customers at i the highest utility choice is making a purchase at an F store in an *SST*, and zero otherwise.

v_{i} One if for the customers at i the highest utility choice is making a purchase at an L store in an *SST* trip and zero otherwise. Defined only if $\exists u_{ik_{t}}^{L} > 0$.

z_{i} One if the utility for customers at i making an *SST* purchase at $k_{t} \in \overline{K}_{i}$ or at an F store j is the same, and zero otherwise. Defined only if $\exists u_{ik_{t}}^{L} > 0$.

y_{ijk}^{qp} One if the highest utility choice for customer i is visiting first a q store at j and then a p store at k in a comparison-shopping or *TST*, and zero otherwise. Defined only for $u_{ijk}^{qp} > 0$.

We now formalize the utilities introduced above. We remark that, as it is customary (see, e.g., Huff 1963), we compute utilities perceived by a consumer for visiting a store and, later, these utilities are used to compute the capture of a store relative to other alternatives.

3.1 Single Stop Trip Expected Utility

The expected utility of a customer at i of making an *SST* to a store p located at site j, is:

$$u_{ij}^p = \alpha \left[r_i - \pi^p - \left(g_{ij} + g_{ji} \right) \right] + (1-\alpha) \left[-\left(g_{ij} + g_{ji} \right) \right]. \tag{1}$$

The first term is the utility of purchasing a product p in an *SST* to site j, weighted by the probability α of making the purchase. The second term is the utility of not making the purchase, having made the trip. If desired, the price could be made dependent on the store.

Note that we assume that consumers do not have an a priori preference among products, which leaves the prices and trip costs as the only drivers of consumers' decisions when choosing between products. However, the probability α can be made to be dependent on the chain. When consumers make only *SSTs*, we assume they choose their highest expected utility store. As locations of the leader's stores are known, to capture consumers at i, the follower F chooses a site j and a price π^F that make $u_{ij}^F > u_{ik}^L$ for all sites k where leader's L stores are located.

3.2 Two-Stop Trip Expected Utility

If the expected utility of visiting two stores is positive, the customers could visit a leader's L store followed by a follower's F store, or the other way around. Once at the second visited store, consumers have complete information on both products, and they choose whether to purchase, and at which store. With a certain probability p_{ijk}^{pq} the purchase is made at the second visited store q and the customer returns home. With a probability $(1 - p_{ijk}^{pq})$, the customer goes back to the first store p, makes the purchase and returns home from there.

The total expected utility that a consumer starting the search at i will perceive, from making the entire trip to a first store p located at j and a second store q located at k, is:

$$u_{ijk}^{pq} = \beta \left[r_i - \left\{ g_{ij} + g_{jk} + p_{ijk}^{pq}(g_{ki} + \pi^q) + \left(1 - p_{ijk}^{pq}\right)\left(g_{kj} + g_{ji} + \pi^p\right) \right\} \right] + (1-\beta) \left[-\left\{ g_{ij} + g_{jk} + g_{ki} \right\} \right] \tag{2}$$

As before, the first term is the utility of making the purchase at some store, while the second term is the utility of visiting both stores and not making any purchase. Once at the second store, the consumer has full information on the quality at both stores. If it happened that $(g_{ki} + \pi^q) = (g_{kj} + g_{ji} + \pi^p)$, i.e., the total costs plus prices of the two alternative actions were the same, p_{ijk}^{pq} would be ½, since we assumed that preferences are evenly distributed. However, when the equality does not hold, we

assume that the customers distribute between the alternatives in an inverse proportion to their total cost, i.e.,

$$p_{ijk}^{pq} = \frac{\left(g_{kj} + g_{ji} + \pi^p\right)}{\left(g_{ki} + \pi^q\right) + \left(g_{kj} + g_{ji} + \pi^p\right)} \tag{3}$$

In terms of purchases at each store, the proportion of the demand originating at i that purchases product q at k is p_{ijk}^{pq}, while $\left(1-p_{ijk}^{pq}\right)$ is the proportion of the demand purchasing product p at store j.

This utility is computed for both possible directions, i.e., visiting first a store q followed by a p store, and the other way around. It may happen that utilities for different trips are the same. Some tie-breaking rules of behavior are required. We assume that, if the utility of an *SST* is equal to the utility of a *TST*, the customer will engage in a *TST*. In addition, if there are trips *i-L-F-i* and *i-F-L-i* with equal utilities, the customer will prefer visiting F first. Preliminary tests have shown that this last rule does not have significant effect on the market capture of both competitors, while allowing a simpler model.

3.3 The Model

The formulation of the *Follower Location model with Comparison-Shopping* is as follows:

$$FLCS : \text{Max } Z_F = \sum_i \alpha a_i\left(v_i + {}^1\!/_2 z_i\right) + \sum_{\substack{i\in I;\, j\in J, \\ k\in K}} \beta a_i p_{ikj}^{LF} y_{ikj}^{LF}$$

$$+ \sum_{\substack{i\in I;\, j\in J, \\ k\in K}} \beta a_i\left(1-p_{ijk}^{FL}\right) y_{ijk}^{FL} \tag{4}$$

$$\text{s.t.} \quad Z_L = \sum_i \alpha a_i\left(v_i + {}^1\!/_2 z_i\right) + \sum_{\substack{i\in I;\, j\in J, \\ k\in K}} \beta a_i\left(1-p_{ikj}^{LF}\right) y_{ikj}^{LF} + \sum_{\substack{i\in I;\, j\in J, \\ k\in K}} \beta a_i p_{ijk}^{FL} y_{ijk}^{FL} \tag{5}$$

$$y_i \le \sum_{\substack{j\in N_i^F \wedge \\ u_{ij}^F > 0}} x_j \quad \forall i\in I \tag{6}$$

$$z_i \le \sum_{j\in N_i^{L=F}} x_j \quad \forall i\in I \tag{7}$$

$$v_i \leq 0 \quad \forall i \in I \,\big|\, u^L_{ik_t} \leq 0 \tag{8}$$

$$y^{FL}_{ijk} \leq x_j \quad \forall i \in I; j \in J, k \in K \tag{9}$$

$$y^{LF}_{ikj} \leq x_j \quad \forall i \in I, k \in K, j \in J \tag{10}$$

$$y_i \geq x_j - \sum_{\substack{r \in N^F_i | \\ u^F_{ir} > u^F_{ij}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{FL}_{irt} \geq u^F_{ij}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{LF}_{itr} \geq u^F_{ij}}} x_r \quad \forall i \in I, \forall j \in N^F_i \,\Big|\, u^F_{ij} > 0 \tag{11}$$

$$v_i \geq 1 - \sum_{\substack{r \in N^F_i | \\ u^F_{ir} > u^L_{ik_t}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{FL}_{irt} \geq u^L_{ik_t}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{LF}_{itr} \geq u^L_{ik_t}}} x_r - z_i \quad \forall i \in I, k_\ell \in \overline{K}_i \,\big|\, u^L_{ik_\ell} > 0 \tag{12}$$

$$z_i \geq x_j - \sum_{\substack{r \in N^F_i | \\ u^F_{ir} > u^F_{ij}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{FL}_{irt} \geq u^F_{ij}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{LF}_{itr} \geq u^F_{ij}}} x_r \quad \forall i \in I, j \in N^{L=F}_i \,\Big|\, u^F_{ij} > 0 \tag{13}$$

$$y^{FL}_{ijk} \geq x_j - \sum_{\substack{r \in N^F_i \cup N^{L=F}_i | \\ u^F_{ir} > u^{FL}_{ijk}}} x_r - \sum_{\substack{k_t \in \overline{K}_i | \\ u^L_{ik_t} > u^{FL}_{ijk}}} 1 - \sum_{\substack{r \in J, t \in K | \\ u^{FL}_{irt} > u^{FL}_{ijk}}} x_r$$

$$- \sum_{\substack{r \in J, t \in K | \\ u^{LF}_{itr} \geq u^{FL}_{ijk}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{FL}_{irt} = u^{FL}_{ijk}; \\ t+r < k+j}} y^{FL}_{irt} \quad \forall i \in I, \forall j \in J, k \in K | u^{FL}_{ijk} > 0 \tag{14}$$

$$y^{LF}_{ikj} \geq x_j - \sum_{\substack{r \in N^F_i \cup N^{L=F}_i | \\ u^F_{ir} > u^{LF}_{ikj}}} x_r - \sum_{\substack{k_t \in \overline{K}_i | \\ u^L_{ik_t} > u^{LF}_{ikj}}} 1 - \sum_{\substack{r \in J, t \in K | \\ u^{FL}_{irt} > u^{LF}_{ikj}}} x_r$$

$$- \sum_{\substack{r \in J, k \in K | \\ u^{LF}_{itr} > u^{LF}_{ikj}}} x_r - \sum_{\substack{r \in J, t \in K | \\ u^{LF}_{itr} = u^{LF}_{ikj}; \\ t+r < k+j}} y^{LF}_{itr} \quad \forall i \in I, \forall j \in J, k \in K | u^{LF}_{ikj} > 0 \tag{15}$$

$$y_i + v_i + z_i + \sum_{j\in J, k\in K | u_{ikj}^{LF}>0} y_{ikj}^{LF} + \sum_{j\in J, k\in K | u_{ijk}^{FL}>0} y_{ijk}^{FL} \leq 1 \quad \forall i\in I \qquad (16)$$

$$\sum_j x_j \leq n_F \qquad (17)$$

$$x_j, y_i, v_i, z_i, y_{ijk}^{FL}, y_{ikj}^{LF} \in \{0, 1\} \quad \forall i\in I, j\in J, k\in K. \text{All } y \text{ and } v \text{ variables defined}$$
$$\text{for corresponding positive utilities} \qquad (18)$$

Objective (4) maximizes the market share of chain F. The first term is the demand of consumers that make *SST*s. The second term includes customers that make *TST*s starting by visiting one of the L stores. The third term are the consumers making *TST*s that start at an F store. There are respectively proportions α and β of consumers that make a purchase in *SST* and *TST* trips. Note that this is not a multiobjective problem and (5) is not an objective, but a constraint, that counts the market share captured by chain L using similar terms to those in the objective. Constraints (6) state that capturing *SST* consumers at i by F is possible only if there are F stores with a utility that is greater than the highest utility L stores. Constraints (7) allow capturing *SST* consumers from i by F stores with the same utility as the highest utility L stores. Later, variable z_i is weighted by ½ in the objective, to represent the capture of only half of the sales. Constraint (8) allows capture of a consumer at i by an L store only if is one of the highest utility L stores. Constraints (9) and (10) ensure that capture by *TST*s is possible if there are *TST*s with positive utility. Constraints (11) to (15) are "highest utility assignment constraints" (Marianov et al. 2018). These constraints correspond to the problem solved by the customer and characterize customers' behavior, i.e., the maximization of utility in a sequential search, and there is one set of these constraints for each possible choice of the customers at i, as explained next. Constraints (11) force the variable y_i to take the value one if an *SST* to an F store is the highest utility choice for i. Note that this only can happen if an F store is located at a site $j\in N_i^F$. If firm F locates such a store (i.e., $x_j = 1$ in the first term on the right-hand side), and if the utility of the trip is positive, the first term on the right-hand side is one. If the remaining terms are zero, the constraint forces $y_i = 1$. If any of the remaining terms on the right-hand side is a non-zero, it will mean that there is a higher utility choice for the customer, and y_i is not forced to take the value one. These higher utility choices are, in the same order they appear in the constraints:

i)   One or more F stores located at points $r\in N_i^F$ provide a higher utility on an *SST*.
ii) One or more *TST*s *i-r-t-i*, visiting an F store located at r and then an L store at $t \in K$, has at least the same utility as the *SST* to j or higher.
iii) One or more *TST*s *i-t-r-i*, visiting an L store located at $t \in K$ and then an F store at r has at least the same utility as the *SST* to j or higher.

Constraints (12) force the variable v_i to take the value one if an *SST* to an *L* store is the highest utility choice for *i*. The first term of these constraints is similar to that in constraints (11), except that the location of all *L* stores is known. The following three terms are the same as in constraints (11), and the last term is one if there are *F* stores with the same utility as the highest utility *L* stores, when customers make *SST*s. Constraint (13) makes variable $z_i = 1$ if *SST*s to some *F* store have the same utility as *SST*s to the highest utility *L* stores. Constraints (14) force $y_{ijk}^{FL} = 1$ if the highest utility choice for customers at *i* is a *TST i* - *F* store - *L* store - *i*, and their explanation is similar to that of the preceding two sets of constraints. Note that, if there are two different trips, say *i-j-k-i* and *i-r-t-i*, with the same utility, picking any of them is indifferent, and the last term of the constraint will choose the one with $t + r < k + j$ to break the tie. Constraints (15) are equivalent to (14), for *TST*s *i* - *L* store - *F* store - *i*. Note that there could be a tie between *TST* utilities from *i-j-k-i* and *i-k-j-i* trips, which would make the model choose a suboptimal solution, due to constraint (16). We avoid that by making the consumers to choose visiting an *L* store first, if such situation occurs. This is enforced by the fifth terms in constraints (14) and (15).

Constraints (16) force the capture of consumer *i* by exactly one facility in an *SST*, or a pair of facilities and the order of their visit in a *TST*. Constraint (17) sets the number of stores to be open by the chain *F*, and constraints (18) defines the domain of the decision variables.

4 Computational Experience

This section reports computational results first on a small, 30-node instance, and then from one thousand 100-node random instances. The number of variables required for the model is $O(4|N| + 2|N|^2|K|)$, the number of constraints is $O(4|N| + 4|N|^2|K| + |N|^2 + 2|N||K| + 1)$, and the problem is NP-hard, as we show next.

Proposition 1 The problem is NP-hard.

Proof: Let *K*, the set of leader sites, be the empty set. The problem reduces to the Maximum Covering Location Problem, which is NP-hard (Megiddo et al. 1983) ∎.

Although the problem is NP-hard, we run the integer-programming model using AMPL and CPLEX 12.8, allowing it to use only one thread in each run. The computer was an HPE Proliant DL360 G9 server with two Intel® Xeon® CPU E5–2630 v4 @ 2.20GHz, 160 GB of RAM, and Debian 9 Operating System.

We first comment on the observability of the parameters, or how these parameters can be estimated in practice. Utilities depend on the product mill price, the reservation price, cost of travel and, in the case of a trip visiting two stores, the probability p_{ijk}^{pq}. The price of the products is observable. There is literature available on how to estimate consumers' reservation price (see Jagpal 2008, Jedidi and Jagpal 2009, Breidert 2006 and references therein). The cost of travel and value of time has also been modeled and empirically measured (see Small 2012; Oregon Department of Transportation 2014; Victoria Transport Policy Institute 2016, and many references therein.) Finally, in this paper we estimate the value of p_{ijk}^{pq} by assuming that if both product are perceived equal, a 50% of the consumers will consider adequate the product in store *L*, and 50%, that in

store *F*. Furthermore, once the consumer compares both products and he is at the second visited store, his decision also depends on the difference between the full costs of making the purchase in both stores. We remark that this rule can be trivially replaced by any other rule that allocates customers to products (e.g., Huff rule), including imbalances in the preferences. The rule can also be estimated using the data available to retailers from their previous experience as well as market studies.

The probabilities α and β are the probabilities of the customers finding and purchasing a suitable product when engaging in single and multiple-stop trips, respectively. The rationale for applying these probabilities is the empirical fact that the number of items of the kind the consumer desires, that are likely to be available during the trip, have an influence on the utility of a trip (Huff 1963). Thus, a *TST*, with more available items, will have a higher utility β. The probability α was set to 0.5, assuming that 50% of the consumers will find acceptable the product they find at the store in a single-stop trip. The value of β covered the range 0.5 to 0.9, reflecting the fact that as consumers are more selective, they value more engaging in comparison-shopping and are less satisfied with visiting a single store. Market studies reflect how selective the customers are. This is in accordance with Huff (1963), although later, he uses the shopping center's square footage of selling space as a proxy for the number of items. Stahl (1982) uses a similar approach to compare markets with different numbers of competing stores.

4.1 Small Instance

We first made runs on a small 30-node instance to show some of the location patterns that result from the assumptions we made. Recall that firm *L*, the leader, locates without foresight, i.e., not considering that there could be another firm entering the market later on. If only the leader is present in the market, consumers' reservation prices, together with the price of the product define how far they are willing to travel to make the purchase at the single provider. The *Maximum Capture Location Model* by Church and ReVelle (1974) can thus be used to solve the leader's problem with no foresight on his part. The entering firm *F*, the follower, wants to enter the market and maximize its market share. If there is no comparison-shopping, the maximum capture (*MAXCAP*) model by ReVelle (1986) solves the problem for the optimal locations of the stores of the entering follower chain, which is what the usual competitive models solve. By setting $\alpha = 1$ and $\beta = 0$ in our model, we reproduce this situation. For a review of competitive models, see Eiselt et al. (2015). In addition, our model includes the context in which there is comparison-shopping, which we model using values of β larger than 0.5.

The 30-node instance is the same as in ReVelle (1986). Without loss of generality, all nodes are both demand points and location candidates. We used $\alpha = 0.5$, and values of β ranging from 0.5 to 0.9. The run time of each instance never exceeded 0.094 s, and the average time was 0.022 s. In Fig. 2, the reservation price is 300, and the prices at both chains are 150. The Figure shows the situation in which both chains locate three stores each, for $\beta = 0.5$, 0.7 and 0.9, in that order.

The Figure shows how, for equal prices, as the probability β of finding a suitable product by visiting two stores increases, the follower tends to co-locate with the leader, as this strategy facilitates the comparison and attracts more customers. Figure 2a shows

the case in which there is no additional gain for visiting two stores over visiting just one, as $\alpha = \beta$. There is no incentive for the follower to co-locate with the leader, as this would mean sharing the market in equal proportions. Rather, the follower locates to obtain a regionally monopolistic market, in which consumers never visit more than one store. As the probability β of the consumers making a purchase increases, enabling comparisons by moving closer to the competitor increases the expected sales, so that the follower tends to co-locate with the leader. Both total and each competitor's market (the number of consumers making a purchase) increase as follows: leader 59, follower 68 ($\beta = 0.5$, Fig. 2a), leader 78, follower 87 ($\beta = 0.7$, Fig. 2b), and leader 169, follower 169 ($\beta = 0.9$, Fig. 2c). Equal demands on both competitors reflect the fact that they are co-located and offer their products at the same price. Note how the number of nodes whose demand does make purchases, increases when moving to the right in Fig. 2.

This behavior is similar for different numbers of stores, as Fig. 3 shows. Note that this case could be representative of the follower desiring to locate stores over a time horizon, in which case the intermediate locations would be robust, i.e., are a subset of the final set of locations.

Table 2 shows the demand for different numbers of stores (#) located by the leader and the follower (*L* and *F*). The third and following columns show the leader and follower demand captured in *SST*s and *TST*s, as well as the total demand captured by each competitor. For $\beta = 0.5$, all captured demand consists of single-stop shoppers, while for $\beta = 0.9$, the captured demand not only significantly increases, but it includes mainly two-stop shoppers.

4.2 Computational Experience with Larger Instances

We run the model for 1000 instances with 100 demand nodes each. The region is a 100×100 square over which the demand is randomly located according to a uniform distribution on each axis, from an integer uniform distribution from 0 to 100. We chose to generate 1000 random instances because these cover many of the possible configurations of demand distribution. The distances are Euclidean, but network or any other practical norms or distances work the same. The competitors could locate one, two or three stores. The reservation price is 200. The probability α is 0.5. The following Tables synthesize the results.
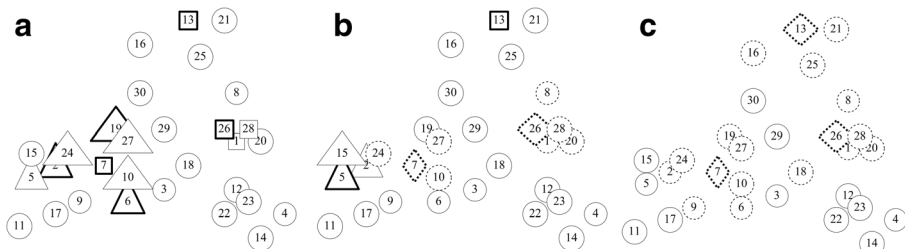
Fig. 2 Leader and follower locate three facilities each. In all cases, $\pi^L = \pi^F = 150$, $\alpha = 0.5$. **a** $\beta = 0.5$, **b** $\beta = 0.7$, **c** $\beta = 0.9$. Circles: customer locations. Dashed border indicates customers visited in *TST*s. Squares: bold for leader location, thin for customers captured by the leader. Triangles: bold for follower location, thin for customers captured by the leader. Diamonds: bold if leader and follower co-locate, thin if both leader and follower capture customers
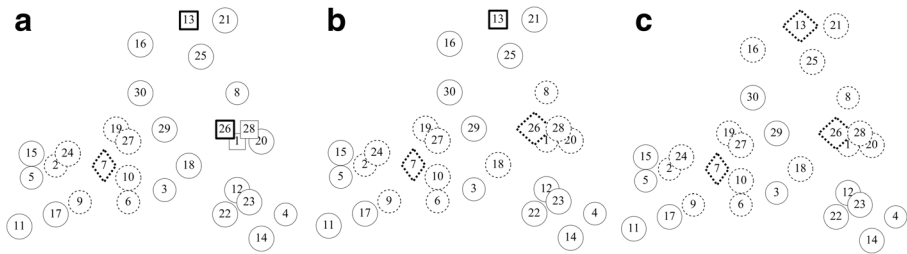
Fig. 3 The leader locates three facilities. In all cases, $\pi^L = \pi^F = 150$, $\alpha = 0.5$, $\beta = 0.9$. **a** follower locates 1 store, **b** follower locates 2 stores, **c** follower locates 3 stores. Same notation as before

Table 3 shows the capture for three leader stores and one, two and three follower stores, when both firms have equal prices. We keep the notation used in Table 2 and the figures correspond to averages over the 1000 instances. The average number of co-located stores is denoted Co-loc. The last column shows the average run time required for solving one instance.

For any number of follower stores, Table 3 shows that as β increases, so does the capture. There is a significant follower advantage in terms of capture for low values of β. As β increases, the follower stores co-locate with the leader's stores. When both have the same number of stores, the capture of both firms becomes the same. Furthermore, with only two stores and for low β, the follower still obtains an amount of demand that is larger than the demand captured by the leader with three stores. As β increases, the two follower stores tend to co-locate with two of the leader's facilities and the third leader's facility gives him an advantage. For only one follower store, its location follows the same behavior: for low β it acts as a regional monopoly but, as β increases, it eventually co-locates with one of the leader's stores.

Table 2 Demand for different numbers of stores and values of β

		$\beta = 0.5$						$\beta = 0.9$					
L #	F #	L SST	F SST	L TST	F TST	L Tot	F Tot	L SST	F SST	L TST	F TST	L Tot	F Tot
1	1	22	28	0	0	22	28	0	0	86	86	86	86
1	2	22	54	0	0	22	54	0	28	86	86	86	113
1	3	22	76	0	0	22	76	0	45	86	86	86	131
2	1	42	26	0	0	42	26	20	0	86	86	106	86
2	2	42	49	0	0	42	49	0	0	135	135	135	135
2	3	42	68	0	0	42	68	0	18	135	135	135	153
3	1	59	26	0	0	59	26	38	0	86	86	123	86
3	2	59	49	0	0	59	49	18	0	135	135	153	135
3	3	59	68	0	0	59	68	0	0	169	169	169	169
3	4	59	85	0	0	59	85	0	18	169	169	169	186

$\pi^L = \pi^F = 150$

Table 4 displays the captured demand for three stores per chain, at equal prices of 120, 150 and 180 and a range of values of β. For β = 0.5, both competitors' stores receive always only *SST* purchases, and the follower has a strong advantage in captured demand (almost double), because the leader located without foresight and the information on leader locations is available for the follower when locating his own stores.

As β increases, the total capture of both chains increasingly includes *TST* demand, the total capture increases for both competitors, and their markets approach each other in size, as co-location increases. When β = 0.9, the total capture is entirely due to *TST* shoppers and both chain markets are practically the same, indicating very strong predominance of co-location. Note that there is a column denoted as "Av min dist", which is the Average Minimum Distance between closest competitors' stores. For each solution, the closest *B* store is found for each *A* store and these distances are averaged. The mean value over the 1000 instances is reported. As β increases, this distance, a proxy for agglomeration, decreases because the follower's stores approach the leader stores. As the price increases, the demand captured by both competitors decreases because the utility of all kinds of trips also decreases. Especially, for very high prices, e.g., 180, agglomeration and *TST*s demand decrease. It is worth noting that although captured demand decreases with price, the profit does not necessarily follow the same trend, and the customers' total utility decreases. When the prices set by both chains are different, the location patterns change, but general trends remain. Co-location is not

Table 3 Average captured demand for 1000 runs on random instances

β	L SST	F SST	L TST	F TST	L Tot	F Tot	Co-loc	Time(s)
One Follower store								
0.5	406	265	0	0	406	265	0	0.03
0.6	404	261	4	4	408	265	0	0.04
0.7	345	173	106	106	451	279	0	0.04
0.8	270	18	345	345	615	362	1	0.05
0.9	265	0	505	505	771	505	1	0.05
Two Follower stores								
0.5	378	491	0	0	378	491	0	0.04
0.6	373	484	7	7	380	491	0	0.04
0.7	272	331	185	185	457	515	1	0.05
0.8	138	54	606	606	743	660	2	0.07
0.9	121	0	925	925	1046	925	2	0.09
Three Follower stores								
0.5	354	694	0	0	354	694	0	0.03
0.6	344	682	12	12	356	694	0	0.04
0.7	223	494	232	232	454	726	1	0.05
0.8	47	133	773	773	821	906	2	0.06
0.9	1	3	1263	1263	1264	1267	3	0.08

Leader: three stores. Follower: one, two and three stores. α = 0.5, $\pi^L = \pi^F = 150$

Table 4 Average captured demand for 1000 runs on random instances

β	L SST	F SST	L TST	F TST	L Tot	F Tot	Co-loc	Av min dist	Time (s)
$\pi^L = \pi^F = 120$									
0.5	637	1117	0	0	637	1117	0	24	0.07
0.6	583	1055	67	67	650	1122	0	23	0.09
0.7	250	531	698	698	948	1228	2	13	0.17
0.8	24	69	1501	1501	1525	1570	3	2	0.26
0.9	1	3	1997	1997	1998	2001	3	1	0.41
$\pi^L = \pi^F = 150$									
0.5	354	694	0	0	354	694	0	20	0.03
0.6	344	682	12	12	356	694	0	20	0.03
0.7	223	494	232	232	454	726	1	15	0.04
0.8	47	133	773	773	821	906	2	4	0.06
0.9	1	3	1263	1263	1264	1267	3	0	0.07
$\pi^L = \pi^F = 180$									
0.5	167	306	0	0	167	306	0	23	0.01
0.6	158	297	10	10	168	307	0	22	0.01
0.7	122	248	69	69	191	317	1	19	0.01
0.8	66	149	208	208	274	357	2	12	0.01
0.9	20	49	400	400	420	449	2	4	0.01

Three stores each chain. $\alpha = 0.5$

always the best strategy, but as β increases the tendency towards agglomeration remains. Table 5 shows this situation.

To analyze l the location patterns in more detail, we made runs varying the value of β in steps of 0.1. Figures 4, 5 and 6 show the changes in capture, both in *SST*s and *TST*s, as well as the co-location and average minimum distances, as β changes.

An analysis of Figs. 4, 5 and 6 shows that, no matter what the price difference between chains, for low values of β the capture is completely due to *SST*s. As β increases, the purchases due to *TST*s increase and those by *SST*s decrease. If the follower's price is low or similar to that of the leader (Figs. 4 and 5), the *SST*s are negligible when β is high. However, when the follower price is higher than the leader's (Fig. 6), there is always some capture by *SST*s at high βs, because two-stop trips are more expensive, and the follower can capture more demand by establishing regional monopolies. It is interesting to note that, for low follower price (Fig. 4) and low β, the follower frequently co-locates with the leader, as this allows sharing high demand areas where the leader is located. As β increases, however, a point is reached at which this strategy is not optimal (β = 0.6, in this case), and the follower changes to a monopoly strategy. Beyond that point, again the follower approaches the leader locations as β increases. For high follower price (Fig. 6) and when β is low, the follower stays away from the leader as its higher price makes him an inferior choice for customers. Rather, he remains as a regional monopoly until the benefit of comparing becomes enough for

Table 5 Average captured demand for 1000 runs on random instances

β	L SST	F SST	L TST	F TST	L Tot	F Tot	Co-loc	Av min dist	Time(s)
$\pi^F = 120$									
0.5	103	1267	0	0	103	1267	0.7	11	0.07
0.6	108	1251	8	10	115	1261	0.5	12	0.07
0.7	100	977	220	270	320	1247	0.2	14	0.18
0.8	18	386	856	1048	874	1434	0.5	7	0.28
0.9	1	80	1388	1695	1389	1775	0.4	6	0.32
$\pi^F = 150$									
0.5	354	694	0	0	354	694	0.0	20	0.03
0.6	344	682	12	12	356	694	0.1	20	0.03
0.7	223	494	232	232	454	726	0.9	15	0.04
0.8	47	133	773	773	821	906	2.4	4	0.06
0.9	1	3	1263	1263	1264	1267	2.9	0	0.07
$\pi^F = 180$									
0.5	422	316	0	0	422	316	0.0	23	0.01
0.6	422	316	0	0	422	316	0.0	23	0.01
0.7	262	237	132	111	394	348	0.9	16	0.01
0.8	186	168	281	236	467	405	1.3	12	0.02
0.9	117	98	491	413	608	511	1.4	8	0.02

Three stores each chain. $\alpha = 0.5$, $\pi^L = 150$. $\pi^F = 120, 150, 180$

breaking this strategy ($\beta = 0.68$). From then on, the *TSTs* increase, as well as the agglomeration as shown by the average minimum distance.
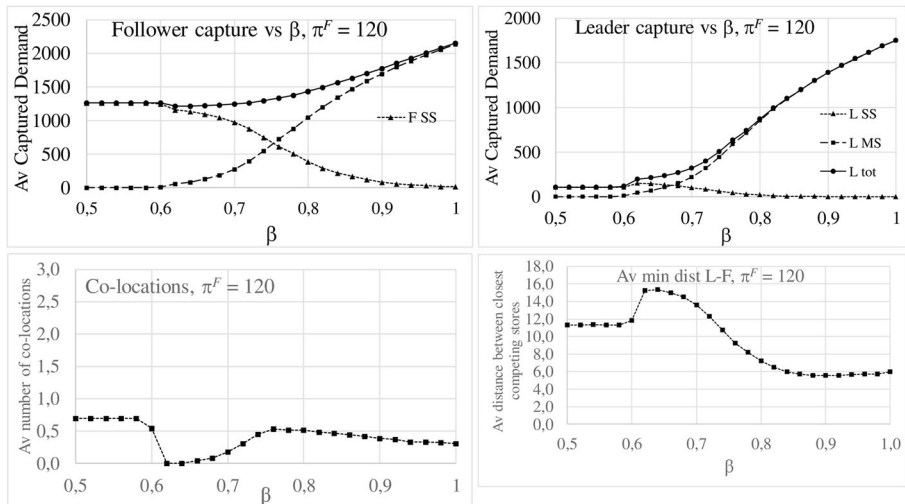

Fig. 4 Above: Follower and leader's capture by *SSTs*, *TSTs*, and total. Below: Average number of co-locations and average minimum distance between closest competitors' stores. $\pi^L = 150$, $\pi^F = 120$, $\alpha = 0.5$. Three stores located by each competitor
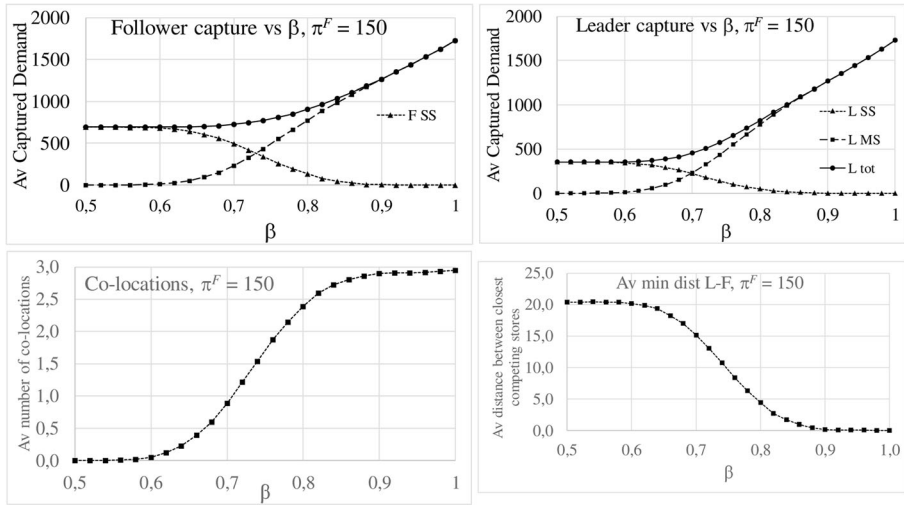
Fig. 5 Above: Follower and leader's capture by *SST*s, *TST*s, and total. Below: Average number of co-locations and average minimum distance between closest competitors' stores. $\pi^L = 150$, $\pi^F = 150$, $\alpha = 0.5$. Three stores located by each competitor

Note that when prices are different, for high values of β, the co-location decreases slightly, because the capture radius of the competitors is different, which makes the follower seek the optimum not necessarily co-locating with the leader. Finally, Table 6 shows the capture for both competitors for different prices and values of β.

Table 6 shows that an increase in the follower price always decreases his capture. When β is small, part of the market lost by the follower by increasing its price, goes to the leader. As β increases however, e.g., 0.8 or above, both firms loss market when the
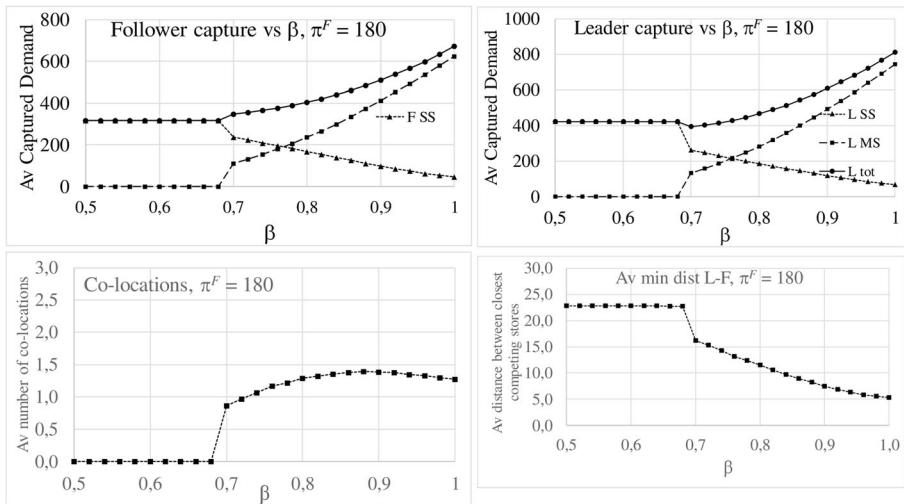

Fig. 6 Above: Follower and leader's capture by *SST*s, *TST*s, and total. Below: Average number of co-locations and average minimum distance between closest competitors' stores. $\pi^L = 150$, $\pi^F = 180$, $\alpha = 0.5$. Three stores located by each competitor

Table 6 Average captured demand for 1000 runs on random instances. Three stores each chain. $\alpha = 0.5$. $\pi^L = 150$

π^F	$\beta = 0.5$		$\beta = 0.6$		$\beta = 0.7$		$\beta = 0.8$		$\beta = 0.9$		$\beta = 1$	
	F Tot	L Tot	F Tot	L Tot	F Tot	L Tot	F Tot	L Tot	F Tot	L Tot	F Tot	L Tot
120	1267	103	1261	115	1247	320	1434	874	1775	1389	2151	1749
130	1048	179	1020	248	1052	323	1219	764	1545	1279	1922	1680
140	865	237	846	307	869	365	999	680	1299	1145	1673	1562
150	694	354	694	356	726	454	906	821	1267	1264	1727	1727
160	556	388	557	386	574	412	656	557	839	822	1112	1158
170	430	409	430	408	455	410	526	512	663	702	877	966
180	316	422	316	422	348	394	405	467	511	608	672	810

follower rises the price. In usual competitive models, an increase in the price of one of the competitors reduces only his own market.

4.3 Computational Experience with a Real Network

Santiago is the capital and the largest city in Chile, with a population of almost 7 million people. We use a 353-node network of Santiago, using its 353 census districts (the union of several census tracks). The location of the nodes of each census district is obtained by allocating each of them to the closest node in the 2212-node main road network of the city, which allows using the road network distances. If two census districts have the same closest road network node, one of them is allocated to its second closest node of the road network.

In the resulting 353-node network, every census districts is connected to its neighbors, through arcs whose length is the shortest distance over the road network. Figure 7 shows the network, on a background representing the Metropolitan Area of Santiago.

A Floyd-Warshall algorithm was applied to the 353-node network, to obtain minimum distances between every pair of nodes. The demand in every node is obtained from the 2002 Chilean National Census. We use a travel cost equal to its distance. Every node is a both a demand point and a potential facility location. Tables 7, 8 and 9 show some of the results, which confirm the insights obtained with the previous runs. We remark that the reported time is the sum of the AMPL plus the CPLEX time. In average, the CPLEX time is between 3 and 4% of the total time.

From Table 7, we see again that as β increases, the locations change from regional monopolies to shared markets (*SST* capture decreases, *TST* capture increases, co-location and average distance between firms both decrease.)

Table 8 shows that, for the parameter values used in these runs, the follower always co-locates as many stores as possible with leader's stores (for five or less stores, all the captured demand is by *TST*s, over five stores, leader and follower have the same amount of demand captured by *TST*s.)

Table 9, finally, shows that naturally, as the follower's price increases, its captured demand decreases. It is interesting to note that for low follower prices, the leader's
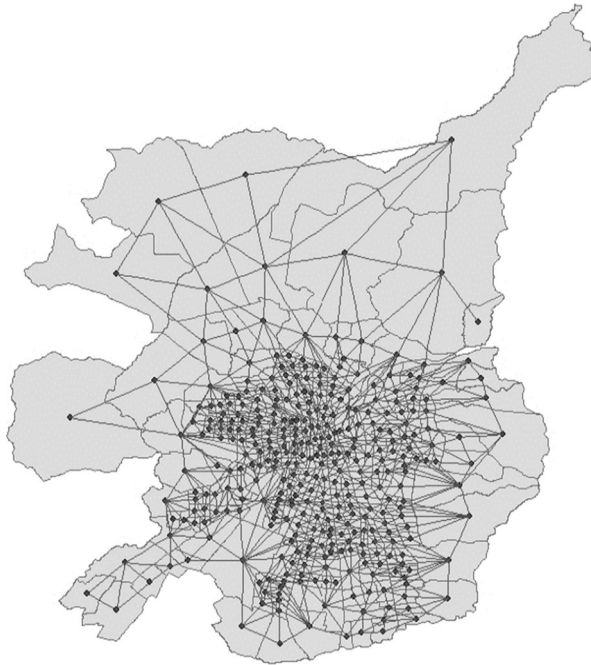
Fig. 7 The 353-node network of the Metropolitan Area of Santiago

market increases. However, as the follower increases its price over the leader's price, both competitors loose market, as they both become less attractive.

5 Conclusions and Future Work

This paper has addressed the competitive location problem, considering that consumers engage in a behavior that is more complex than those studied so far in the location

Table 7 Average captured demand for the 353-node network of Santiago

β	L SST	F SST	L TST	F TST	L Tot	F Tot	Co-loc	Av min dist	Time(s)
0.50	287,348	419,394	0	0	287,348	419,394	0	6884	4.38
0.55	287,348	419,394	0	0	287,348	419,394	0	6884	6.14
0.60	287,348	419,394	0	0	287,348	419,394	0	6884	6.32
0.65	227,431	335,839	86,076	86,076	313,506	421,914	1	6307	6.51
0.70	84,701	175,275	277,160	277,160	361,861	452,435	3	3709	7.81
0.75	33,013	93,832	413,812	413,812	446,824	507,643	4	428	7.56
0.80	0	0	611,951	611,951	611,951	611,951	5	0	7.57
0.85	0	0	695,365	695,365	695,365	695,365	5	0	8.03
0.90	0	0	842,657	842,657	842,657	842,657	5	0	8.98

Five stores each chain. $\alpha = 0.5$, $r = 15,000$, $\pi^F = \pi^L = 5000$

Table 8 Captured demand for the 353-node network of Santiago

L #	F #	L SST	F SST	L TST	F TST	L Tot	F Tot	Time(s)
5	1	255,947	0	182,284	182,284	438,230	182,284	9.32
5	2	204,259	0	361,368	361,368	565,626	361,368	8.94
5	3	133,406	0	536,768	536,768	670,174	536,768	8.92
5	4	59,593	0	708,788	708,788	768,381	708,788	8.92
5	5	0	0	842,657	842,657	842,657	842,657	8.90
5	6	0	89,377	842,657	842,657	842,657	932,034	8.89
5	7	0	170,820	842,657	842,657	842,657	1,013,477	8.86
5	8	0	250,710	842,657	842,657	842,657	1,093,367	8.89
5	9	0	319,964	842,657	842,657	842,657	1,162,620	8.88
5	10	0	386,840	842,657	842,657	842,657	1,229,496	8.89

The leader locates five stores, the follower locates between one and ten stores. $\alpha = 0.5$, $\beta = 0.9$, $r = 15,000$, $\pi^F = \pi^L = 5000$

literature. In particular, instead of assuming that consumers make single trips to a store where they buy a good, we include comparison-shopping, the behavior according to which consumers visit stores belonging to different retail chains, before deciding whether making a purchase, and where. In our case, two chains or firms, leader and follower, selling each one product or brand of products, locate their stores sequentially, and consumers decide their trips once the stores are located. The products sold at the chains are imperfect mutual substitutes, i.e., they differ in secondary characteristics as

Table 9 Captured demand for the 353-node network of Santiago

π^F	L SST	F SST	L TST	F TST	L Tot	F Tot	Co-loc	Av min dist	Time(s)
1000	–	–	691,741	1,409,767	691,741	1,409,767	–	933	15.95
2000	–	–	750,965	1,195,931	750,965	1,195,931	–	844	14.03
3000	–	–	735,968	1,011,892	735,968	1,011,892	1	908	12.57
4000	–	–	759,273	874,357	759,273	874,357	2	700	11.18
5000	–	–	842,657	842,657	842,657	842,657	5	–	10.16
6000	–	–	707,790	626,845	707,790	626,845	1	704	9.27
7000	–	–	690,736	545,251	690,736	545,251	1	794	8.63
8000	–	–	681,233	477,814	681,233	477,814	1	775	8.00
9000	–	–	643,720	407,747	643,720	407,747	1	595	7.49
10,000	7494	–	630,498	365,882	637,992	365,882	1	595	7.06
11,000	59,593	50,336	535,713	289,913	595,306	340,249	2	1757	6.76
12,000	73,813	50,336	480,330	239,788	554,143	290,124	1	1915	8.22
13,000	133,406	97,362	344,167	160,143	477,573	257,505	–	4708	8.02
14,000	151,773	97,362	331,372	143,572	483,145	240,934	1	4550	7.82

The leader and the follower locate five stores each. $\alpha = 0.5$, $\beta = 0.9$, $r = 15,000$, $\pi^L = 5000$

color, texture, etc. Consumers are aware of the prices charged by the chains. We assume that consumers make a first decision at their point of origin, regarding what store or pair of stores to visit. This decision does not change along the way. When visiting one store, a consumer can either make the purchase or not. If visiting two stores, after comparing the products, the consumer decides first whether purchasing the product or not, and then, in what of the visited stores to purchase it.

The decision on what store or pair of stores to visit is made based on the utilities perceived by the customer a priori, which depend on the reservation price, the prices at the stores, the total distance to be traveled and the probabilities of finding a suitable product. The trip chosen in the first stage is the one with the highest utility. In the second stage, the decision is made based on products' features.

We propose a model for the consumer comparison strategy and an optimization model for the follower location problem. We run a small instance from the literature, showing that the comparison-shopping behavior and its consideration by the follower, increases the likelihood of agglomeration or co-location. We also run the model on one thousand 100-node instances generated randomly, whose statistics confirm the findings and show that the model is solvable in short run times.

The main results include

- an indication that an increase of β (the probability of finding a suitable product by visiting two stores) increases agglomeration,
- an indication that an increase of β increases the demand captured by both competitors,
- other things being equal, the follower captures significantly more demand than the leader (a result dubbed the "first entry paradox" by Ghosh and Buchanan 1988),
- an increase of the price level of the follower by 50% resulted in between a 66% and a 75% of decrease of capture,
- for markets in which there is no strong drive for comparing products, an increase in the price of the follower decreases his market and increases or keeps unchanged the leader's market, and
- for strong comparison-shopping-oriented customers, an increase in the follower's price level significantly decreases the market capture of *both* firms.

Furthermore, if consumers are not selective (the probabilities of purchasing a suitable product by visiting one store or two stores are close to each other), firms locate their stores as regional monopolies. Such a store, say k, could increase the price of its product as long as the utility of consumers in the neighborhood for purchasing at k is higher than the utility of going to the closest competitor's store. On the other hand, if customers are more selective, they must compare, and as a result, stores tend to co-locate in spite of the fact that they have to share the market. In this case, any price increase comes with a proportional loss of customers. However, they do so because the probability of obtaining a purchase is higher. This differs from the situation in which there is multipurpose shopping (Marianov et al. 2018) in which agglomeration or clustering is due to the fact that consumer surpluses when buying two products are higher, and consumers have an incentive to travel farther away to get both products. In synthesis, there are two opposing forces when locating near to the competitor: the competition becomes stronger, as firms must share the market. However, by facilitating

comparison-shopping, the agglomerated stores attract more customers. The hope is that this market increase counterweights the effect of sharing.

The model we present is a one-level version of a two-level problem, in which the entering chain solves the first level (the location of stores, conditional to the location of existing stores), and the consumers the second level (the choice of the store at which the purchase is made. This model provides a tool to investigate the clustering behavior observed when there is comparison-shopping, and to quantify the effects of consumers' selectivity or preferences.

The model includes a rule on how the consumers choose the purchase once the comparison is done. This rule can be easily be replaced by other rules (e.g., Huff), and preferences or biases towards one or the other chain can also be included trivially.

The knowledge obtained from solving the follower problem allows us to address the leader problem in a duopoly as future research. Note that the leader problem is a game with three players: the leader, the follower, and the consumers. The formulation, as it is now, is not generalizable to the case of more than two chains, as it would require more assumptions on customer behavior. However, the information obtained from the results on the duopoly, allows extracting the basics and later, addressing more complicated cases. We intend to address this problem in the future. There are more potential research directions to be explored in the context of multi-stop shopping. A first one is related to different customers' searching strategies, e.g., allowing changes of mind along the trip, or taking into account the fact that the decisions have random components. Natural extensions include more than two products and other discrete choice rules.

References

Aras N, Küçükaydın H (2017) Bilevel models on the competitive facility location problem. In: Mallozzi L et al (eds) Spatial interaction models. Springer Optimization and Its Applications 118

Arentze TA, Oppewal H, Timmermans HJP (2005) A multipurpose shopping trip model to assess retail agglomeration effects. J Mark Res 42:109–115

Berglund PG, Kwon C (2014) Solving a location problem of a Stackelberg firm competing with Cournot-Nash firms. Netw Spat Econ 14:117–132

Breidert C (2006) Estimation of willingness-to-pay: theory, measurement, application. Deutscher Universitäts Verlag, Wiesbaden. 135 pages

Brown S (1989) Retail location theory: the legacy of Harold Hotelling. J Retail 65(4):450–470

Chamberlin E (1933) The theory of monopolistic competition. Harvard University Press, Cambridge

Church R, ReVelle C (1974) The maximal covering location problem. Papers of the Regional Science Association 32(1):101–118

D'Aspremont C, Gabszewicz JJ, Thisse JF (1979) On Hotelling's 'stability in competition'. Econometrica 47: 1145–1150

Dellaert BGC, Arentze TA, Bierlaire M, Borgers AWJ, Timmermans HJP (1998) Investigating consumers' tendency to combine multiple shopping purposes and destinations. J Mark Res 35:177–188

Drezner Z (1982) Competitive location strategies for two facilities. Reg Sci Urban Econ 12:485–493

Drezner T (2014) A review of competitive facility location in the plane. Logist Res 7:114. https://doi.org/10.1007/s12159-014-0114-z. last retrieved 7/21/2019

Drezner T, Drezner Z (1996) Competitive facilities: market share and location with random utility. J Reg Sci 36:1–15

Drezner T, Drezner Z, Eiselt HA (1996) Consistent and inconsistent rules in competitive facility choice. J Oper Res Soc 47:1494–1503

Eaton BC, Lipsey RG (1975) The principle of minimum differentiation reconsidered: some new developments in the theory of spatial competition. Rev Econ Stud 42(1):27–49

Eaton BC, Lipsey RG (1979) Comparison-shopping and the clustering of homogeneous firms. J Reg Sci 19(4):421–435

Eaton BC, Lipsey RG (1982) An economic theory of central places. Econ J 92(365):56–72

Eiselt HA (2011) Equilibria in competitive location models. In: Eiselt HA, Marianov V (eds) Foundations of location analysis. Chapter 7. Springer Science + Business Media, New York, pp 139–162

Eiselt HA, Marianov V (eds) (2015) Applications of location analysis. Vol 232 in the international series in operations research & management science. Springer-Verlag, Cham. 437p

Eiselt HA, Marianov V, Drezner T (2015) Competitive location models. In: Laporte G, Nickel S, Saldanha-da-Gama F (eds) Location science. Chapter 14. Springer-Verlag, New York

Fernández J, Pelegrín B, Plastria F, Tóth B (2007) Solving a Huff-like competitive location and design model for profit maximization in the plane. Eur J Oper Res 179:1274–1287

Ghosh A, Buchanan B (1988) Multiple outlets in a duopoly: a first entry paradox. Geogr Anal 20:111–121

Ghosh A, McLafferty S (1984) A model of consumer propensity for multipurpose shopping. Geogr Anal 16(3):244–249

Guo WC, Lai FC (2014) Spatial competition with quadratic transport costs and one online firm. Ann Reg Sci 52(1):309–324

Hakimi SL (1983) On locating new facilities in a competitive environment. Eur J Oper Res 12:29–35

Hodgson MJ (1978) Toward more realistic allocation in location-allocation models: an interaction approach. Environ Plan A 10:1273–1285

Hodgson MJ (1990) A flow-capturing location-allocation model. Geogr Anal 22:270–279

Hotelling H (1929) Stability in competition. Econ J 39:41–57

Huff DL (1963) A probabilistic analysis of shopping center trade areas. Land Econ 39:81–90

Jagpal S (2008) Fusion for profit: how marketing and finance can work together to create value. Oxford University Press 664 pages

Jedidi K, Jagpal S (2009) Willingness-to-pay: measurement and managerial implications. In: Rao V (ed) Handbook of pricing research in marketing. Chapter 2. Edward Elgar Publishing Ltd., Cheltenham. 616 pages

Kress D, Pesch E (2012) Sequential competitive location on networks. Eur J Oper Res 217:483–499

Kress D, Pesch E (2016) Competitive location and pricing on networks with random utilities. Netw Spat Econ 16:837–863

Laporte G, Nickel S, Saldanha-da-Gama F (eds) (2015) Location science. Springer-Verlag, New York

Lerner AP, Singer HW (1937) Some notes on duopoly and spatial competition. J Polit Econ 45(2):145–186

Marianov V, Eiselt HA (2016) On agglomeration in competitive location models. Ann Oper Res 246:31–55

Marianov V, Ríos M, Icaza MJ (2008) Facility location for market capture when users rank facilities by travel and waiting times. Eur J Oper Res 191(1):32–44

Marianov V, Eiselt HA, Lüer-Villagra A (2018) Effects of multipurpose shopping trips on retail store location in a duopoly. Eur J Oper Res 269(2):782–792

McLafferty S, Ghosh A (1987) Optimal location and allocation with multipurpose shopping. In: Ghosh A, Rushton G (eds) Spatial analysis and location-allocation models. Von Nostrand Reinhold, New York, pp 55–75

Megiddo N, Zemel E, Louis Hakimi S (1983) The maximum coverage location problem. SIAM Journal on Algebraic Discrete Methods 4(2):253–261

Miralinaghi M, Keskin BB, Lou Y, Roshandeh M (2017) Capacitated refueling station location problem with traffic deviations over multiple time periods. Netw Spat Econ 17:129–151

Mulligan GF (1987) Consumer travel behavior: extensions of a multipurpose shopping model. Geogr Anal 19(4):364–375

Nagy G, Salhi S (2007) Location-routing: issues, models and methods. Eur J Oper Res 177:649–672

O'Kelly M (1981) A model of the demand for retail facilities, incorporating multistop, multipurpose trips. Geogr Anal 13(2):134–148

O'Kelly M (1983) Impacts of multistop, multipurpose trips on retail distributions. Urban Geogr 4(2):173–190

Okabe A, Suzuki A (1987) Stability of spatial competition for a large number of firms on a bounded two-dimensional space. Environment and Planning A: Economy and Space 19:1067–1082

Oregon Department of Transportation - Economic & Financial Analysis Unit (2014) The Value of Travel-Time: Estimates of the Hourly Value of Time for Vehicles in Oregon 2013. Available in http://library.state.or.us/repository/2009/200902121450374/2013.pdf, last retrieved on 7/21/2019

Pelegrín B, Fernández P, García MD (2018) Computation of multi-facility location Nash equilibria on a network under quantity competition. Netw Spat Econ 18:999–1017

Prescott EC, Visscher M (1977) Sequential location among firms with foresight. Bell J Econ 8(2):378–393

ReVelle C (1986) The maximum capture or "sphere of influence" location problem: Hotelling revisited on a network. J Reg Sci 26(2):343–358

Santos-Peñate DR, Campos-Rodríguez CM, Moreno-Pérez JA (2019) A kernel search Matheuristic to solve the discrete leader-follower location problem. Networks and Spatial Economics Online. https://doi.org/10.1007/s11067-019-09472-7

Small K (2012) Valuation of travel time. Econ Transp 1(1):2–14

Stahl K (1982) Differentiated products, consumer search, and locational oligopoly. J Ind Econ 31:97–113

Thill JC (1982) Spatial duopolisic competition with multipurpose and multistop shopping. Ann Reg Sci 26: 287–304

Victoria Transport Policy Institute (2016) Transportation cost and benefit analysis: techniques, estimates and implications [second edition]. Available in http://www.vtpi.org/tca/, last retrieved on 7/21/2019

von Stackelberg H (1943) Grundlagen der theoretischen Volkswirtschaftslehre (translated as The theory of the market economy). W. Hodge & Co., Ltd., London. 1952

Wolinsky A (1983) Retail trade concentration due to consumers' imperfect information. Bell J Econ 14(1): 275–282