Advances in methods to support store location and design decisions

Hunneman, Auke

Chapter 2

Location and Design of Multiple Stores in a Competitive Environment: A Mixed Integer Linear Programming Approach[1]

2.1 Introduction

One of the fundamental decisions a retailer must make is the location of new stores. The importance and complexity of such decisions have increased, due to the rapid growth of multistore networks (Ghosh, McLafferty, and Craig 1995). Today's retail industries are characterized by high levels of concentration. In the United States for example, retailers with four or more stores account for more than half of total retail business (Pal and Sarkar 2002). Location decisions are highly complicated for networks of stores operated by a single firm, because they require the systematic evaluation of the impact of each store on the entire network and the consideration of interactions among stores (Ghosh, McLafferty, and Craig 1995). From a consumer perspective, distance to the store is not the only factor driving store choice; differences in designs across stores play an important role as well. Consumers who choose among various competitors offering a more or less identical products do not automatically visit the nearest store (see, e.g., Briesch, Chintagunta, and Fox 2009; Chernev and Hamilton 2009); they also consider, for example, multipurpose traveling, comparison shopping, and differences in prices, goods, assortments, and images among stores (for an overview, see Pan and Zinkhan 2006). Consequently,

[1] This chapter is based on a working paper with the same title, co-authored by Matthijs Streutker.

configuring multistore networks is a nontrivial task that requires the simultaneous optimization of store location and design in a competitive environment. In this chapter, we propose a model to support these decisions for a chain of stores. The model contributes to extant literature because it can be solved as a mixed integer linear programming (MILP) model, which is very flexible in its specification and can be solved effectively for instances with realistic sizes.

The remainder of this chapter is organized as follows: First, we summarize the relevant literature (section 2.2). In section 2.3, we formulate our general competitive facility location and design problem model. Some extensions of the model, which explain the behavior of consumers or the preferences of the company in more detail, also appear in this section. In section 2.4, we illustrate the model with a case study that describes the location and design of luxury leisure clubs in the Rotterdam area of the Netherlands. We also provide intuitions about model behavior through a comparison with the classical full capture model[2] and a sensitivity analysis of the findings for the store size sensitivity parameter. The concluding remarks and directions for further research are in section 2.5.

2.2 Literature Overview

This chapter considers the competitive facility location problem: Finding the optimal sites for several retail outlets operating in a competitive environment. In principle, it is a multi-agent optimization problem, where the agents are the retailer, the consumers, and competitors. All agents have their own (conflicting) goals (e.g., profit maximization for the retailer and competitors, utility maximization for consumers) and make their own decisions (e.g., store location/size, store choice). This complicated problem cannot be solved easily, so simplifying assumptions about the behavior of consumers and competitors are necessary. Many models

[2] This model also has been referred to as the maximum capture (MaxCap) model. We prefer the term full capture model.

assume competitors are already present in the market, and their locations and other characteristics remain fixed for the time period considered. We restrict ourselves to this static setting as well (for an overview of similar competitive facility location models, see Plastria 2001). Thus, there are no competitor decisions, and the only two actors are the retailer and consumers.

The question then becomes how to incorporate consumer decisions into models that optimize the locations and design of new stores. Consumers usually concentrate in a finite set of points, such that each point represents a geographical entity, such as a region, city, or zip code. Patronage likely depends on utility functions that measure the attractiveness of a store to a consumer, generally based on travel distance and perhaps other attraction factors, such as store size and (perceptions of) the store's price level. In full capture models, as initiated by Revelle (2006), each consumer patronizes the most attractive store for him or her. Therefore, all consumers concentrated in a demand point visit the same store. This so-called deterministic patronizing behavior represents a strong assumption; alternatively, in spatial interaction or gravity models, all stores are visited by the group of consumers represented by a particular demand point in proportion to their relative attractiveness. These models, introduced by Huff (1964), feature probabilistic patronizing behavior, and Santos-Peñate, Suárez-Vega, and Dorta-González (2007) and Serra and Colomé (2001) describe the differences in the two classes. In particular, both store choice, given attractiveness to consumers, and the attraction function itself can be specified in several ways. In optimization models, the most widely exploited formulations are additive and multiplicative, for which the attractiveness of a store to a consumer depends on several factors. As the names suggest, a multiplicative attraction function uses the product of weighted factors, whereas the additive version calculates the sum. Discussions of these options appear in Drezner, Drezner, and Eiselt (1996) and Drezner and Eiselt (2002).

Important developments on the retailer side of the problem include the less spatially restricted location decisions and endogenous determination of the number of stores to be located. Originally, competitive facility location problems were formulated in discrete space, such that the number of candidate sites were determined in advance, and the store(s) were located in a subset of these sites. The network approach is more flexible; stores can be located at a node or along an edge of a predefined network (e.g., Suárez-Vega, Santos-Peñate, and Dorta-González (2004). More recent models propose a continuous planar space, such that the facilities can appear anywhere on the plane defined by the demand points. Fernández et al. (2007) offer a spatial interaction continuous model; Plastria and Carrizosa (2004) outline a full capture version. Introducing design as a decision variable in competitive facility location models also changed the model set up, because the number of facilities became endogeneous instead of exogeneous. Without design questions, the user specifies the number of facilities, with the objective of maximizing revenues. Without any costs involved, the endogeneous determination of the optimal number of facilities results in a facility at each potential location. However, because design induces costs, the objective changes to maximizing profits for multiple facilities, determined endogeneously.

The competitive facility location and design problem (CFLDP) thus is difficult to solve, regardless of the specific formulations—deterministic or probabilistic, discrete, network, or planar. Until recently, models could evaluate only the performance implications of a (finite) number of location and design alternatives. Enumeration schemes can be effective for the location and design of a single facility when the number of possible locations is limited and the optimal size for a certain location can be computed effectively (Eiselt and Laporte 2006). Furthermore, recent effective optimization methods support specific formulations of planar single-facility location and design problems, such as the maxcovering-minquantile problem reformulation by Plastria and Carrizosa (2004), disk covering reformulation by Zhang (2001), Weiszfeld-like algorithm and interval branch-and-bound algorithm by Fernández et al. (2007), and universal evolutionary global optimizer by Redondo et al. (2009). Multiple facilities also complicate the planar case; to the best of our knowledge, Tóth et al. (2009) are the only ones to consider this problem for two facilities. They apply the interval branch-and-bound algorithm, but when demand points increase in number to 50–100, this algorithm takes many hours to solve. Moreover, discrete competitive facility models also are complicated by the introduction of design variables. Therefore, many authors consider design fixed (e.g., Benati 1999; Benati and Hansen 2002; Colomé, Lourenço, and Serra 2003), and contributions that pursue "true" design and location optimization for multiple stores are rare. Zhang and Rushton (2008) suggest a genetic algorithm to solve the problem, but they cannot provide numerical results. Suárez-Vega, Santos-Peñate, and Dorta-González (2004) solve a network CFLDP for multiple facilities by combining a global search procedure with three combinatorial heuristics. Unfortunately, the optimality of their solutions is not clear. The tangent-line approximation algorithm by Aboolian, Berman, and Krass (2007), with finite design alternatives, provides this information; specifically, they find optimal solutions for instances with 60–80 nodes and two design alternatives within 90 minutes.

We propose in turn a full capture model for locating and designing multiple stores in a competitive environment, which can be solved as a mixed integer linear programming (MILP) model. This approach not only allows for a continuum of design alternatives but also can include all kinds of constraints, which explain either the behavior of consumers or the preferences of the retailer. We have sufficient knowledge about solving MILP models, such that instances of considerable sizes can be solved effectively. Moreover, this well-known paradigm is widespread in decision-oriented applications, which eases model adoption. Nevertheless, we break in some sense with current research traditions in CFLDP, because we model deterministic, instead of probabilistic patronizing behavior. Furthermore, the attraction function is additive instead of multiplicative, and our location space is discrete instead of planar. A probabilistic model for store patronage seems realistic, but in the end, what matters most is whether the model provides good location and design suggestions. The objective of CFLDP models is not to forecast store patronage, which instead is the next step in a location study that uses detailed (regression) models and can include many (explanatory) variables. Store patronage obviously differs between probabilistic and deterministic models. For example, in a location-only study, Serra and Colomé (2001) show that with uncertainty about true patronage behavior, the deterministic model provides the most robust solutions. In addition, regarding the specification of the attraction function, classic utility theory suggests a multiplicative function is more appealing than an additive one. However, Drezner, Drezner, and Eiselt (1996) argue that the additive rule is more consistent, because with the additive rule, preferences do not change en route. Although planar CFLDP thus is an intriguing mathematical problem, we believe that for practical implementations, the discrete model suffices. We therefore discuss the specifications of a general CFLDP model in more detail.

2.3 Model

Consider a retailer that wants to open several new stores in a certain region. This retailer needs to know (1) how many stores to open, (2) where to locate these stores, and (3) the optimal size for each store. These new stores will compete for consumers with the stores of competitors that already operate in this region. We assume competition is static and known in advance. The retailer aims to maximize profit with these store openings; we assume it does not have any incumbent stores in the region yet, though this assumption easily can be relaxed. Demand and potential store locations may coincide and are represented by points. We use travel distance to calculate distances between each pair of points, which indicates the actual distances consumers travel by car to get from one point to another.

A demand point represents a populated area, such as a village or city district. Demand for products sold at these stores should be inelastic, such that the volume of demand does not depend on product quality. Consumers exhibit deterministic patronage behavior and choose the store most attractive to them. We model the attractiveness of a store with an additive utility function that consists of two explanatory variables: travel distance and store size.

Store points can be population areas as well, but other strategic locations in the region to be specified by the retailer. Each store point is as a candidate site for a new store, and they differ in costs. Some store points are occupied by stores of competitors; no store point can support more than one store, so they comprise potential locations for new stores and locations of competitors. Relaxing this condition would lead to a novelty-oriented tie-breaking rule (for a discussion on tie-breaking rules in full capture models see, e.g., Plastria 2001), because building a slightly bigger store at the same point as a competitor means capturing all the clientele from the competitor's store.

Notation

For the formulation of the model we use the following notation:

Sets

P set of all points, with the following subsets:

P^d set of demand points,

P^p set of possible store points, and

P^c set of competitor store points.

Note that $P^p \cap P^c = \varnothing$ by the one store per store point assumption.

Variables

x_j = 1 if a store opens at possible store point j, and 0 otherwise, $j \in P^p$; for

$j \in P^c$, x_j is a parameter with $x_j = 1$.

s_j = size of the store at possible store point j, $j \in P^p$; for $j \in P^c$, s_j is a parameter.

x_{ij} = 1 if a consumer of demand point i shops at store point j, and 0 otherwise, so $i \in P^d$, $j \in P^p \cup P^c$.

z_{ij} = size of the store where a consumer of demand point i shops, $i \in P^d$, and $j \in P^p \cup P^c$.

y_j = revenues in store point j, $j \in P^p \cup P^c$.

Parameters

c_j = variable cost per unit size for store at store point j, $j \in P^p$.

s = minimum size of a store.

d_{ij} = distance between demand point i and store point j, $i \in P^d$, and $j \in P^p \cup P^c$.

β = size sensitivity parameter.

A_i = decisive attraction for consumers at demand point i, $i \in P^d$.

δ_i = demand of consumers in demand point i, $i \in P^d$.

M = a large number (maximum realistic size of a store).

Formulation

Before we can formulate the objective and constraints, we need to know consumers' attractions to the incumbent stores before the store opens. First, for each consumer, attraction to each store is calculated using the auxiliary parameter A_{ij}:

$$A_{ij} = -d_{ij} + \beta s_j,$$

for all $i \in P^d$ and $j \in P^c$. Note that for $j \in P^c$, s_j is a parameter. The attraction a consumer has toward a certain store depends linearly on the distance required to travel to it and its size. Note that β relates store size to travel distance. In most CFLDP models, it is the other way around: distance gets discounted by design attractiveness. We choose this formulation instead, because with it we can easily turn off the size effect, which facilitates the comparison of the results with the location-only full capture model when we set β to 0. Second, the consumer visits the store with the highest attraction. Following Plastria and Carrizosa (2004), we call this attraction the decisive attraction of a consumer, given by

$$A_i = \max_{j \in P^c} A_{ij}, \tag{2.1}$$

The location constraints are as follows:

$$x_j \in \{0,1\}, \text{ and} \tag{2.2}$$

$$s_j \geq s x_j, \tag{2.3}$$

for all $j \in P^p$. Constraint 2.2 requires that at a certain point j, either a new store opens ($x_j = 1$) or not ($x_j = 0$). We also specify a minimum size for new stores. If a store opens at a certain location ($x_j = 1$) its size must be at least s (Constraint 2.3).

The locations of new stores together with the configuration of incumbent stores determine shopping possibilities, which submit to the following constraints:

$$x_{ij} \in \{0,1\}, \tag{2.4}$$

$$x_{ij} \leq x_j, \tag{2.5}$$

$$z_{ij} \leq s_j, \text{ and} \tag{2.6}$$

$$z_{ij} \leq M x_{ij}, \tag{2.7}$$

for all $i \in P^d$ and $j \in P^p \cup P^c$. Constraint 2.4 requires that a consumer spends his or her entire budget in a certain store or nothing at all. A consumer can only shop at a certain location if there is a store located there, as noted in Constraint 2.5. The size of the store where a consumer shops cannot be larger than the actual size of the shop at that location (Constraint 2.6), and store choice and size choice are linked (Constraint 2.7).

Next, the following constraints ensure that consumers are allocated to stores only if they are at least as attractive as the currently most attractive store:

$$\sum_{j \in P^p \cup P^c} x_{ij} = 1, \text{ and} \tag{2.8}$$

$$\sum_{j \in P^p \cup P^c} \left(-d_{ij} x_{ij} + \beta z_j \right) \geq A_i, \tag{2.9}$$

for all $i \in P^d$. Equation 2.8 requires that each consumer is assigned to exactly one store location, and Equation 2.9 ensures that consumers can be assigned to a new store only if their attraction for it is greater than that for the old situation (without new stores).

The choices of the consumers determine the revenues generated at each store location, given by:

$$y_j = \sum_{i \in P^d} \delta_i x_{ij},$$

for all $j \in P^p \cup P^c$.

The objective function the retailer wants to maximize is:

$$\sum_{j \in P^p} \left(y_j - c_j s_j \right).$$

Profits are equal to total chain revenues minus costs, which depend on store sizes.

Remarks

This model maximizes the retailer's profits through a systematic evaluation of a number of potential locations in a particular market; in this process, it allocates consumers to stores. The stores to which a consumer can be allocated are determined by Equation 2.9. It is easy to see that a consumer is not necessarily allocated to the store for which he or she has the highest attraction. However, if consumers visit the stores they like most, chain profits remain the same. This assertion follows from two observations: First, the retailer is only concerned about whether a consumer visits a store of its chain or a competitor. From a chain perspective, it is not important which store in the network the consumer visits.

Second, if a consumer moves to a new store, the store for which he or she has the highest attraction must belong to this particular retail chain as well. A consumer can only be allocated to a new store if the attraction level for this store is higher than the decisive attraction. Therefore, the store the consumer likes most has an attraction level higher than the decisive attraction and must be one of the new stores.

The parameter of interest is β, or the size sensitivity parameter, and this quantity (provided it is unequal to 0), differentiates the proposed model from the location-only full capture model. The beta measures indicate the relative importance that consumers attach to store size compared with travel distance when deciding where to shop. Low values of β indicate that consumers find store size less important than travel distance; in the extreme case that β is 0, travel distance is the only factor considered. Therefore, the proposed model becomes nearly a classical location-only full capture model, except that costs enter the newly proposed models. In contrast, high values of β make distance less important. Ultimately, if beta is very large, the retailer opens only one store that is slightly larger than the largest competitive store at the location with the lowest costs. In the empirical application, we vary the value of β to investigate how it affects the sizes and locations of the stores to be opened. It is difficult to judge in advance how the objective function changes in response to different values of β, which appears on both the left- and right-hand sides of the model (Equations 2.9 and 2.1).

Extensions

Because we use the MILP framework to solve the CFLDP, the preceding general model can be extended in various ways; we discuss two. The first extension pertains to the level of behavioral detail of the utility functions, and the second deals with possible decision-making preferences from a retailer's perspective.

Consumer Behavior. In the current model, travel distance and store size explain the attraction of consumers to each store, though it also is possible to include other variables in the utility functions, such as product assortment, prices, and opening times. Evidence also indicates that consumers differ in their responses to marketing activities (Pauwels et al. 2011). Retailers seek to exploit these differences by tailoring their marketing mix to the specific needs of local consumer groups. From a marketing perspective, it therefore is reasonable to consider multiple consumer types, which can be defined based on sociodemographic variables and differ in the numerical values of the attraction parameters. In the empirical study, we elaborate on and illustrate this extension.

In the previous model, we implicitly assumed that consumers were willing to travel any distance to satisfy their demand. Prior literature instead indicates that some consumers only will travel a limited distance for certain products, known as the range of a product (e.g., Craig, Ghosh, and McLafferty 1984). If there is no store within this distance, consumers switch to substitutes. If we thus include a maximum travel distance in the model, we must add the following constraint:

$$\sum_{j \in P^p \cup P^c} d_{ij} x_{ij} \leq d_i,$$

in which the maximum distance a consumer from demand area i wants to travel is denoted by d_i. Therefore, a lost sales variable also must be added for each demand area.

Decision Making Preferences. The current model specification determines the optimal number of new stores endogeneously, though it is possible to set a bound on this number, whether an exact number or a minimum or maximum value. If we denote the number of new stores to be located by n, we add the following constraint to the general model to ensure that exactly n new stores are located:

$$\sum_{j \in P^p} x_j = n. \tag{2.10}$$

Profit maximization may be one of multiple objectives a retailer pursues in a particular market. A retailer also might consider it important to limit the market shares of competitors, which we can model by a multiple goal objective. If we define w^p and w^c as weights for the sales of new stores and competitor stores, respectively, the multiple goal objective function takes the following form:

$$-\sum_{j \in P^p} c_j s_j + w^p \sum_{j \in P^p} y_j - w^c \sum_{j \in P^c} y_j .$$

2.4 Empirical Application

We test the model using semi-realistic data from an international chain of luxury leisure clubs that wants to extend its business in the Netherlands. The issues faced by its management are typical for the expansion of multistore networks, which boil down to the determination of the appropriate number of new stores to open in a particular market, their optimal locations, and ideal sizes.

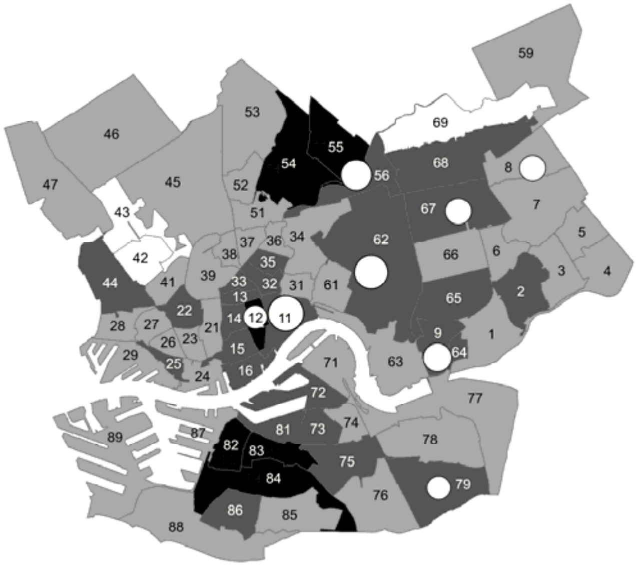
We compare the outcomes of the proposed model with those of several benchmark models. First, we investigate whether the optimal locations of the proposed model differ from those of a location-only full capture model. We also contrast our findings with those of an extended full capture model that includes costs. Second, we evaluate the performance of the proposed model, in which club locations and sizes are optimized simultaneously against an approach that first locates clubs using the extended full capture model and then identifies the optimal sizes for each club. Third, we assess the sensitivity of the location and size results to different values of the size sensitivity parameter through a simulation study. This section ends with a discussion of the results of one of the model extensions, namely, the inclusion of multiple consumer types.

2.4.1 Data

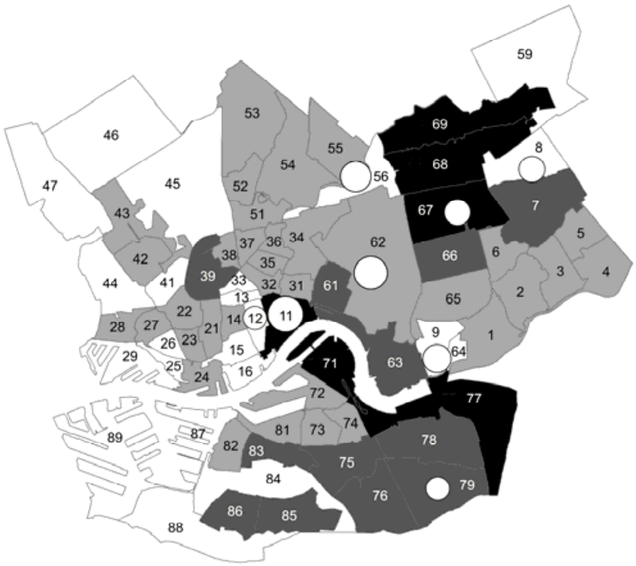
We assume that the management of this health club chain has already decided on the region in which to locate new club(s), namely, the Rotterdam area in the Netherlands, which includes Rotterdam and the neighboring village Capelle aan de IJssel. Decision makers typically separate the identification of the most appropriate region from the choice of particular sites within this region (Levy and Weitz 2004; Mendes and Themido 2004). We subdivide the Rotterdam area into four-digit zip codes, which results in 74 polygons (see Figure 2.1), or "store points" in our model. (In the figures, we omit the first two digits of each zip code, because only the last two digits are unique for all zip codes considered.) We further assume that no clubs of this particular chain are present in the Rotterdam area yet (an assumption that can be easily relaxed). The locations and sizes of competitors appear in Table 2.1 and Figure 2.1. We also use the zip codes as demand points, and for each combination, we obtain travel distances in miles from an online route planner (http://www.mapquest.com).

Table 2.1: Competitors

Location	08	09	11	12	56	62	67	79
Size (m^2)	1000	1200	2000	800	1700	1800	1000	800



[Costs in Euros per year per m^2 with categories, from light to dark: < 200, 200–300, 300–400, and > 400.]



[Demand in thousand Euros per year with categories, from light to dark: < 25, 25–50, 50–75, and > 75.]

Figure 2.1: Costs and demand for each zip code and location of competitors.

Notes: Dot diameter denotes relative size of club.

The first step is to determine the spatial distribution of demand for fitness, which is not readily observed. We use the Statline database of Statistics Netherlands (http://statline.cbs.nl) to determine the number of inhabitants of each zip code. Because the chain offers high-quality facilities, people with higher incomes are its target customers.[3] Of this group, 25% are likely to join a fitness club, and the average membership fee is approximately 40 Euros (IHRSA 2006). Unfortunately, we do not have data to estimate the size sensitivity parameter but instead solve the model for a range of values for this parameter, measured in miles/1000m^2.

We assume the costs of each leisure club consist of those associated with employees (wages) and housing (rent). We obtained information on rental prices in the Rotterdam area from the Realnext Web site (http://www.realnext.com), on which brokers provide information about the (regional) supply of commercial real estate. We then calculated average rental prices for commercial real estate in each zip code. Costs of employees are based on an analysis of company reports from this chain (available in the AMADEUS database[4]). We divided total employee costs by the chain's overall floor space and obtained an amount of 120 euros per m^2.

2.4.2 Results

The model was written in the algebraic modeling language Mosel and solved by Xpress 19 of Dash Optimization (http://www.dashoptimization.com) on a 2.33 GHz Intel Core2Duo machine with 2 GB of RAM. The automatic tuning facilities helped find good values for the many MILP settings.

Results of the Benchmark Models

In this section, we first discuss the results of the location-only full capture model. The standard full capture model assumes that consumers visit the nearest club. This model specification can be obtained from the model proposed in the previous

[3]Consistent with the definition of Statistics Netherlands, we define higher incomes as incomes above the 40 percent point of the Dutch income distribution

[4]http://www.bvdep.com/en/amadeus.html.

section by setting costs and the size sensitivity parameter to 0: $c_j = 0$ for all $j \in P^p$, $\beta = 0$, and we include the number of stores constraint from Equation 2.10. For a number of stores ranging from 1 to 6, the outcomes are listed in Table 2.2 (all solutions found within one second). The first column shows the number of stores considered, whereas the subsequent dots denote the zip codes for new clubs. The last column expresses the overall chain revenues for each configuration (the objective function in this setting). The dominant locations for new clubs are 15 and 35: Location 15 is slightly to the west of the most western competitor and obtains revenues from consumers living in the western part of the region, whereas location 35 serves the northern part of the market.

Table 2.2: Full capture results: Store locations and corresponding revenues
Notes: A dot indicates the presence of a store in a particular zip code.

# Stores	Zip code							Revenues
	02	15	35	68	73	78	81	
1		●						668,158
2			●				●	1,148,470
3		●	●				●	1,389,666
4	●	●	●		●			1,611,632
5	●	●	●	●			●	1,824,786
6	●	●	●	●	●	●		1,953,950

We also consider the extended full capture model, which includes costs, so the c_i's are restored to their original values, but distance remains the only factor driving consumers' store choice. Because size does not affect consumers' choices of a particular club but does affect costs, we only locate clubs of a minimum size of 1000 m^2. As the results of this model in Table 2.3 show, none of the locations found for the basic full capture model is part of the solution for the new model. Frequently occuring zip codes are 21, 36, and 74. Zip codes 21 and 36 are the low-cost equivalents of zip codes 15 and 35 that dominate the standard full capture model. Location 74 is a centrally located, low-cost zip code in the southern part of the region in which competition is low. Opening three clubs at these locations is optimal from a chain profit perspective.

Table 2.3: Results for full capture with costs
Notes: A dot indicates the presence of a store in a particular store.

# Stores	Zip code							Profit
	06	21	36	39	69	74	77	
1		●						363,505
2				●		●		557,733
3		●	●			●		598,756
4		●	●		●	●		577,910
5	●	●	●		●	●		501,765
6	●	●	●		●	●	●	363,168

Results of the Proposed Model

This section discusses the results of the proposed model, which assumes that consumers decide to visit a particular leisure club according to its size and the distance required to travel to the club. Therefore, the size sensitivity parameter (β), measuring consumers' trade-off between club size and travel distance, must be different from 0. To investigate the effect of this parameter on model outcomes, we vary it and solve the model for each value of the parameter. Specifically, we varied β from 0 to 0.6 in steps of 0.001, so we solved the model 600 times for different values of β . Table 2.4 shows the optimal locations and sizes of clubs for several focal values; the two righmost columns contain the profits and solution times for each situation. In Figure 2.2, we display these results graphically. The horizontal axis of this graph indicates different values of the size sensitivity parameter, whereas the cumulative size of all clubs opened appears on the vertical axis. The shaded areas represent the zip codes in which new clubs are opened and their respective sizes in each situation. These findings confirm the "extreme" solutions we anticipated in section 2.3. For small values of β , the results are in line with the distance-only full capture with costs model, whereas high values result in a single large club. Moreover, the results indicate nonmonotonic behavior of the objective value as a function of β, which means that in the interval 0–0.6, profits both increase and decrease over the range of β values. This result emerges because β is part of both the left- and right-hand sides of the model. Regarding club locations, we find that zip codes 21, 36, and 74 are still dominant, just as in the full capture with costs model. However, for β s in the range of 0.15–0.55, location 69 also comes into play. It offers low costs and is located in the high-demand northeastern part of the region. For β s in [0.35,0.5], zip code 21 gets replaced by 23, which is slightly less expensive.

Table 2.4: Results for location and design model

β	21	23	36	61	69	74	Profit	Sol. time
0.00	1000	0	1000	0	0	1000	598,757	< 1
0.05	1000	0	1000	0	0	1000	591,264	1
0.10	1000	0	1000	0	0	1000	587,102	2
0.15	1000	0	1000	0	1600	1000	590,872	4
0.20	1000	0	1000	0	1450	1000	584,116	7
0.25	1000	0	1000	0	1360	1000	596,760	12
0.30	1000	0	1000	0	1300	1000	605,189	18
0.35	0	1029	1000	0	1257	1000	612,133	27
0.40	0	1000	1000	0	1225	1000	624,236	37
0.45	0	1000	1044	0	1200	1000	617,645	75
0.50	0	1020	1140	0	1180	1000	593,422	202
0.55	1000	0	1218	0	1164	1000	574,750	787
0.60	0	0	0	6667	0	0	638,151	375

Results of Two-Stage Procedure

Because the optimal club locations of the distance-only full capture model with costs appear in the solutions to our proposed model as well, we might question whether it is necessary to perform an analysis that identifies optimal club locations and designs simultaneously. Seemingly, the less time-demanding two-stage procedure suffices, because we can first solve the full capture with costs model to identify optimal club locations and then determine the best club sizes for each location. Therefore, to investigate how this two-stage model performs compared with our approach, we conduct both analyses and plot the results in Figure 2.3.

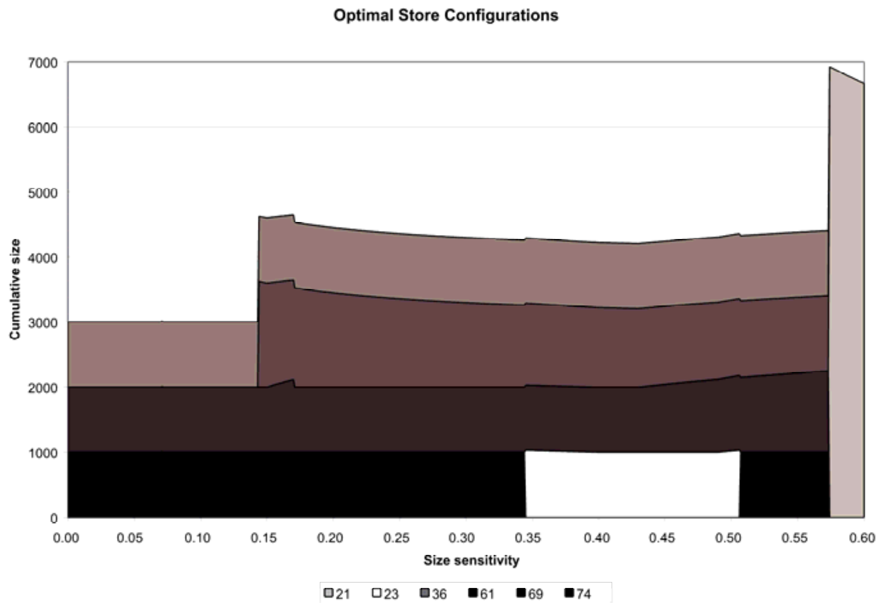


Figure 2.2: Cumulative sizes in the proposed model.

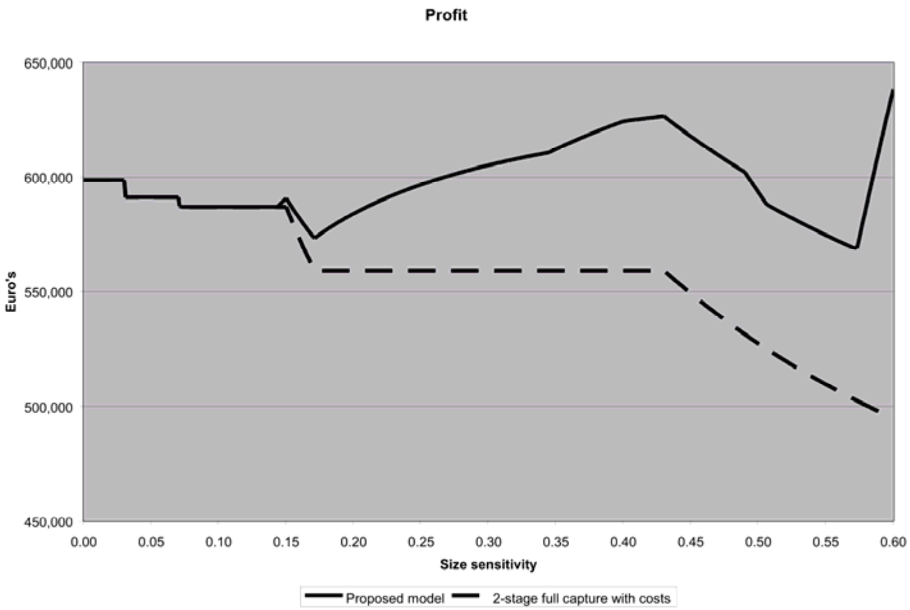


Figure 2.3: Profits of the proposed and two-stage full capture with costs models.

The overall chain profits of the CFLDP model and two-stage model refer to β s in the range [0,0.60]. For low values of β, profits do not differ between the two models, because they result in equal store configurations. Medium values of β result in a CFLDP model profit level approximately 10% higher than that of the two-stage model. This considerable difference becomes much larger for high beta values.

Multiple Consumer Types

The proposed CFLDP model can be extended to include multiple consumer types. In this section, we therefore investigate the effects of distinguishing between two consumer groups: young professionals and families. Consistent with the customer profiles developed by this particular fitness chain, we define these groups as people aged 25 to 35 years (young professionals) and people 35 years and older (families). These two groups also differ in the value of their size sensitivity parameter. Thus, we split all sets, parameters, variables, and equations for the consumers into two subsets. Demand for fitness is determined using demographic data (i.e., population sizes of the two groups) from Statistics Netherlands.

Table 2.5: Results for two consumer types

β^{yp}	β^{f}	21	23	36	39	69	74	Profit	Sol. time
0.300	0.300	1000	0	1000	0	1300	1000	605,189	59
0.354	0.275	1000	0	1000	0	1327	1000	601,358	60
0.408	0.250	0	1000	1000	0	1360	1000	598,750	66
0.462	0.225	0	1000	1070	0	1400	1000	577,223	72
0.516	0.200	1013	0	1000	0	1450	1000	547,568	98
0.571	0.175	1000	0	1246	0	1514	1000	535,243	87
0.625	0.150	1000	0	1000	0	1600	1000	521,273	90
0.679	0.125	1013	0	1000	0	0	1000	513,937	90
0.733	0.100	0	1000	1072	0	0	1000	486,313	111
0.787	0.075	0	1000	1136	0	0	1000	471,812	129
0.841	0.050	0	0	0	1000	0	1000	459,992	189
0.895	0.025	0	0	0	1000	0	1039	449,977	209
0.949	0.000	0	1000	0	0	0	1094	449,564	183

In Table 2.5, β^{yp} and β^{f} are the size sensitivity parameters for young professionals and families, respectively. We assume that families are less likely to travel long distances if they visit a leisure club, and young professionals pay more attention to the size of the club they visit. The first row of Table 2.5 provides the results for a situation in which there is no difference in club size sensitivity between the two groups; in each subsequent row, heterogeneity increases. We lower the size sensitivity of the families in small steps while increasing the sensitivity of young professionals. In each case, the value of β^{yp} is such that the average size sensitivity of the whole consumer population equals the starting value of 0.3.

The results of this extended model are very similar to those of the general model with no consumer heterogeneity. As we show in Table 2.5, the same zip codes are part of the solutions to both models. The only exception is zip code 39, which is a more central location than zip codes 21 and 23, so clubs in these latter zip codes are replaced by that in zip code 39 when travel distance becomes more important to families. Demand by families in each zip code is generally higher than that of young professionals; the configuration of new stores thus closely follows the spatial distribution of demand for this group if β^{f} is very small. Another important observation is that the club located in zip code 69 gets bigger when consumers are more hetereogeneous in their preferences. Therefore, the club opened at this location effectively lures young professionals away from (smaller) competing clubs nearby. With maximum consumer heterogeneity, only two clubs with (almost) minimum sizes open in zip codes 23 and 74, which represent areas with low competition and costs, where family demand for fitness is relatively high. Another notable observation is that the chain's profits decrease monotonically as consumer heterogeneity increases, perhaps because heterogeneous consumer preferences mean each store can only satisfy the needs of one consumer group: either large (young professionals) or close (families), which by definition results in less sales to the other group.

2.5 Conclusions and Further Research

The key objective of this chapter has been to provide a modeling approach to the competitive facility location and design (CFDLP) problem, when the goal is to determine simultaneously the optimal locations and sizes of multiple stores in a competitive environment. The proposed model contributes to (static) competitive facility literature in several important ways. First, it enables us to generate an optimal solution rather than evaluate a fixed number of predefined store configurations. Second, our model offers a continuum of design alternatives and can include all kinds of constraints. Therefore, a new store can be of any size (instead of a limited number of size alternatives), and the behavior of consumers and/or company preferences can be represented realistically. Mathematically, we formulate the model as a mixed integer linear programming (MILP) model, a well-established paradigm that eases the model adoption and implementation. Moreover, knowledge about solving MILP models is sufficient, such that we can solve instances of considerable sizes effectively.

In the empirical study, we illustrate this model by applying it to a realistic case study of the location of new health clubs in the Rotterdam area. The solution of the full model with consumers who are not very sensitive to store size is similar to that of the full capture (with costs) model, whereas for extremely size-sensitive consumers, a single large club is optimal (from the company perspective). Moreover, profits depend nonmonotonically on the value of the store size sensitivity parameter. We also evaluate whether a model with sequential location and design, which demands less time to compute, can generate similar solutions. Consistent with Tóth et al. (2009), we find that the difference in profits between sequential and simultaneous approaches is considerable, depending on the store size sensitivity parameter. Finally, we consider a multiple consumer type extension.

The reasonable solution times for problems of considerable sizes underline the model's usefulness for supporting location and design decisions in practice. Retailers that want to expand their operations to a particular region with several stores opening simultaneously can use the model to find an optimal number of stores for a particular area, identify suitable locations for each store, and determine their optimal designs. In a later stage, a more detailed analysis performed for each site can evaluate its attractiveness and sales potential.

We acknowledge several limitations of our study that suggest directions for further research. In particular, an important and significant challenge will be to find an empirically validated estimate of the store size sensitivity parameter, which we believe is worth the effort. Validating this parameter would make it possible to use it for simulation and prediction purposes. Because we do not know the exact attitudes of individual consumers toward travel distance and store size, we assume the parameters measuring the effects of these variables on store attraction vary across consumers and follow some general distribution (e.g., normal). Another area for research therefore would include constraints that limit the amount of cannibalization—that is, the loss of sales in a particular store due to the overlap of its trade area with that of another store in the same chain. Such modeling additions are crucial if a franchisor wants to add new stores in an area in which it already operates. Sales through franchise systems have grown rapidly in the United States in recent decades (Kalnins 2004; Lafontaine and Shaw 1999), so we consider this extention highly relevant.