22.1. RETAIL LOCATIONAL ANALYSIS

22.1.1. Spatial retail location

The demand by consumers for retail goods and services is a function of the attributes of the commodity, household income, and other factors such as home ownership status. For example, a home improvement store is likely to target a market with a housing stock that has lots of possibilities for repair, upgrades, and remodels. Both home-owning and renting populations might yield adequate density of demand, but the effective demand for goods and services by homeowners is much more likely to be attractive to this particular service. An electrical supplier would have a demand for services also, but for that business it could be some combination of over-the-counter sales (light fixtures) and more substantial electrical equipment sold to contractors and builders. A business with a traditional central market place location (in an older mixed use inner city neighborhood for example) might conceivably want to branch out its locations to catch the growth in the suburbs and even the outlying communities in the hinterland of that main market. In fact there are so many different ways to imagine the dynamics of retail site location that there is a real need for a general purpose simulation tool that might enable the estimation of the merit of various growth proposals (Baker, 2000; Munroe, 2001). In all these cases, it is important to have an accurate estimate of the spatial distribution of effective demand as arising out of a combination of preferences, and disposable income.
Central place theory has long held that there is a hierarchy of goods, from frequently demanded inexpensive items to high-end expensive goods. There is both a higher spatial frequency of demand for (and provision of) the so-called lower-order goods, and a corresponding scarcity on the landscape of higher-order goods. Thus for every Mercedes or Lexus dealership in the city there might be numerous Ford and Toyota dealerships. The higher the order of goods provided, one assumes that there is a wider market scope required to provide sufficient demand to cover the operating costs of the business (the so-called threshold). Similarly, the higher-order goods, because of their relative scarcity on the landscape, require longer trip lengths; the break even calculation for the retailer is whether the spatial extent of the market required to cover costs is matched by a corresponding willingness of consumers to travel to the center for the goods (see the classic study by Berry (1967)).

Inexpensive ‘low order goods’ are sometimes sold in combinations with higher priced items from superstores that do not necessarily have a small range: they can in fact be attractive over a large distance, provided the assortment and price point allows the large agglomerated retailer to undercut the smaller more widely dispersed providers of retail services. This formula is used by Wal-Mart or other ‘big box’ retailers; they have a large assortment of goods, and price points that are competitive, and locations that in themselves act as a magnet for spatial interaction (Munroe, 2001). Customers travel to stores and therefore the spatial interaction of the purchasers must be recognized as an important behavioral factor. The retail and trade area service location problem requires knowledge of what customers want, where they are located, and that they have income that covers the price and market segment of the goods. In common with many levels of retail operation, many of the most successful chains study a massive amount of geo-demographic profile data that enables a rich portrait of customers and consumer behavior to inform the merchandize and market planning of their operations.

22.1.2. Consumer demand and behavior

Measuring the total income and pool of expenditure is accomplished by combining a count of households by geo-demographic cluster (e.g. Claritas PRIZM, MapInfo PSYTE, ESRI Tapestry, AGS Mosaic etc.), the index value for each group (m), the penetration rate and some index of average per household expenditure. To calculate the potential pool of expenditure for the zone i and commodity ‘c’ a formula such as the following might be used:

\[ O_{ic} = \sum_{\text{over all groups}} N_{im} y_{mc} \]

where \( O_{ic} \) is the demand in zone i for commodity c, \( N_{im} \) is the number of group m households in zone i, and \( y_{mc} \) is an expenditure rate per household cluster m on commodity c.

From this aggregate, demand shares allocated to a particular store have two components: on the one hand the share is smaller for the more distant competitive stores (holding other factors constant), but additionally it is felt that the demand for a store increases with the accessibility of that origin zone to any shopping destination. Zonal accessibility, and hence aggregate demand is a function of where the stores are located, and so unlike conventional location models, we should not treat demand as an exogenous factor (O’Kelly, 1999). While it would be naive to say ‘build it and they will come,’ it is certainly reasonable to think that the provision of retail services can induce demand for that service that would otherwise be allocated to other discretionary uses. Some insight and market-based intelligence is needed to capture the correct demand parameters and
sensitivity to locational access. The basic accessibility of each zone can be predefined, and the demand in the immediate area of a new potential store opening can increase, as a result of improved accessibility. One practical estimation approach that can be effective is to have a variety of alternative sources of judgment (like the so-called ‘Delphi’ method, and a variant of the judgmental methods advocated for several years by Seldin (1995)) with perhaps one figure coming from an estimate of per capita expenditure and saturation, one coming from pro forma estimates of expected sales per square foot and yet another estimate coming from an experienced local commercial real estate professional. The best model is likely to use some aspect of these data as controls on the judgmental estimate. In other words, no analyst will simply apply a sales per square foot figure to an arbitrary new built store and say that the expected sales are a product of the coefficient and the store size. Much more likely is an analysis that takes the current sales situation of the competition into account and then projects how much of these existing sales can be captured by the new proposed location. Even more significant is recent research that has shown that whatever the general relationship between the variables, the strong likelihood of spatial variability in such a relationship ought to be taken into account (Fotheringham et al., 2002; Rust and Donthu, 1995). Thus, if a cross-sectional regression analysis provided evidence of a coefficient of say $30 weekly sales per square foot, then a spatially varying parameter might trend significantly with location given the socio-economic patchwork of the city, by analogy with a similar argument in the context of house prices (Fotheringham et al., 2002).

One way to make operational estimates is through spatial interaction models. These models are the topic of this chapter, which covers a variety of models largely inspired by several years experience as both an applied and as a theoretical exploration of retail sales and interaction.

### 22.1.3. The role for models

Among the most basic general questions for spatial interaction modelers are the following: Where do the customers come from? What are the spatial interaction patterns governing the distribution of distance and attraction parameters? What is the probability that a customer at \( i \) patronizes a store at \( j \)? Conditional upon the location of \( i \) what is the probability of being a customer of destination \( j \)?

**Example**

A grocery store has an upper income target consumer. Their research shows that these are very likely to be loyal customers of the produce and fresh foods departments (which in turn are highly profitable assuming that stock can be turned over rapidly to avoid waste/spoilage). In seeking new store locations, where are there sufficient pockets of un-met demand among this target population? In the analysis of existing ‘own-brand’ stores, which may or may not be currently well-located vis-à-vis the standard customer profile, is there any need to modify their type of store to meet consumer needs?

These kinds of questions can be answered with spatial models. Before we get into the details of how to formulate and apply such a model, it may be very helpful to get a preview of some of the uses to which a model might be put. One common usage is in ‘impact analysis.’ With a fitted model, purporting to describe the allocation of consumers to demand centers, we can estimate impact on remaining stores if a branch is to be closed, or indeed if we open a new one. Both of these changes have impacts across the system of stores, but of course the ‘first law of geography’ (Tobler, 1970) which holds that things are more highly interrelated when they are in close proximity, leads us to expect that the impacts are greatest on the centers and competitors closest to the site of the change.

Other uses for fitted models in locational analysis include assessment of the desirability of overhauling various stores or facilities.
The applied retail analyst is often asked to estimate the impact of a change on the expected sales of the store: thus having a model which has as its ‘independent variables’ some measures which can be adjusted to reflect the new attraction of the store can be useful to estimate the change in the retail trade area, expected sales, and so on. By estimating a well-fitted model to these data, we replace the specifics of the data instance with a model that has ‘effects’ – these are systematic influences on the trends in the levels of spatial interaction, and are likely to include roles for distance and retail attraction (typical basic variables in SI models – see Guy (1991)). In more elaborate settings these models can also include many other independent variables (see especially the Multiplicative Competitive Interaction Models MCI – Nakanishi and Cooper (1974)). Once these models are fitted, the analyst can then dial in various changes in the driver variables, and assuming that the model is reasonably robust to changes in these data, the impact of the changes on the expected sales and interaction levels can be determined.

Models can also be used as a tool in assessment of complex strategic questions. For example, a chain that is considering opening a new branch in a growing suburb might be faced with the question of whether to keep an existing older store in a nearby location. The question is then one of strategy: do stores A and B together make a better combined profitable solution than the option of closing B, presumably giving A an even greater new opening sales level, but possibly exposing the chain to the risk that a competitor might take the abandoned site? Not only does the decision hinge on the aggregate sales of the various combinations of open stores but it also must answer questions about the probable impact of competitors. Retailers engage in strategic behavior, and open or close locations as part of a system of decisions; such analyses often include issues of pre-emption and blocking competition, and beating competitors to the punch in new areas of expansion (Ghosh and Craig, 1983).

Models are also useful in assessing ongoing measures of store performance and may be used in this way as an early warning of emerging shifts in the market. Assuming that the chain can collect data throughout its system on the performance of each store, and some appropriately calculated variables to describe the stores site and situation the analyst can embark on the kind of ‘analog’ assessment made popular in the early days of quantitative analysis. This method, in its modern guise, uses the stores sales (as a dependent variable) and a selection of measurements of the trade area characteristics, and develops a multiple regression model to assess the expected (or predicted) sales vs. the actual observed levels. Fundamental to this operation is a meaningful definition of trade area: it makes no sense to include measurement of the ‘attributes’ of areas far away from a store, if indeed it is known that few if any shoppers come from that area. So, in other words, the measurement of the trade area of the store becomes the first and most important operation. There is no hard and fast consensus on how to define a trade area, and much more will be said on this matter later. For now, suffice it to say that the trade area could be objectively defined as an area within 5 min drive time of the store. That leads to a computation of the demand that exists within that area, and that could be one of the independent variables. (Clearly if we use more sophisticated definitions of trade areas, the trade area demand calculation would have to be re-computed.)

Independent variables collected for all the stores are saved as the columns of a table. GIS is especially helpful to calculate features of trade area and give a quantitative descriptive nature of a trade area. The dependent variable is the actual sales performance: there is often a challenge obtaining these data (i.e., weekly sales) in academic research; but it is important to know how these data would be used in an applied case study.

Basing store closing decisions on this kind of result places a lot of faith in the fitted model (and so the importance of regression
diagnostics, measures of goodness of fit, and significance levels on the estimated coefficients). What looks like an under-performer may not actually be an instant argument for store closure: for instance a store is projected to draw $400,000 per week, but actual sales come in at $350,000 (i.e., $50,000 below the regression line). While all might agree that it could be doing better (i.e., it is performing below potential) there may be good reasons that the store has not yet reached its full potential. It might be under attack from particularly aggressive competitors, be poorly managed, or it might be built ‘over-sized’ in anticipation of further population growth in the area. The store may suffer from a depressed regional economy, and so a chain may consider shutting it down or attempting to reinvigorate the system by investing a lot of money into the regional advertising campaign. It all boils down to choices, and these choices are best informed by analytic models.

The ease of obtaining a good fit to the model will clearly vary across sectors. Department store sales volumes are notoriously difficult to predict, in that their aggregate sales volume is a combination of the various heterogeneous departments, and the extent of competition for specific categories in these stores could very well vary in an unsystematic way across locations. On the other hand goodness-of-fit for convenience related stores such as grocery chains are likely to be quite acceptable, in that there are a few predictable variables that are very highly correlated with the aggregate performance of the store. For example, the store’s size, its population base, and the immediate competitive environment undoubtedly account for the bulk of the store-to-store variation in sales levels. Thus, it is expected that the coefficient of store size, and population and competition will be significant, and that the resulting fitted model will have a strong R-square.5 Refinements to the model to include regional dummy variables and other more precise measures of target market demand (through surrogates such as parking studies, or traffic flow) are likely to help to improve the model.

Some sectors lend themselves readily to analysis by multivariate regression models (grocery stores) but others require a different approach. If a shoe store, book store, or branch of a chain of clothing stores is typically located in shopping centers, then the analyst might use the center as a surrogate for the size of the market in which the individual store is located (see also Prendergast et al., 1998). Similarly if a chain of this type is planning to enter a new regional market, it could very well limit its attention to the shopping centers. This type of work is useful because it is frequently necessary to manage thousands of location across many areas/regions.

It is hard to get information on gross sales (what is also called ‘turnover’ in the British literature) in academic case studies, though practitioners and consultants can of course gain access to their client’s data as part of their confidentiality agreement. Many of the ideas in this chapter have been framed as a result of real world experience. In practice, one has access to lots of data; in theory one might have to learn these techniques in a data vacuum, recognizing that the proprietary data would become available to a consultant doing these analyses for a private sector client. This perhaps accounts for the lack of precision in the published literature – a lot of literature in retailing location modeling is quite imprecise mathematically – and the details are often not published in a way that makes verification and validation easy.

### 22.1.4. Consumer choice

The probabilistic assignment of consumers to retail destinations can be formulated as a production constrained spatial interaction model:

$$P_{ij} = A_i O_i W_j \exp(bC_{ij}).$$

Such models calculate the probability that a user at a specific origin location will select
one from a number of available alternative attractive destinations. If these destinations are shopping centers, for example, the attraction of those centers can be represented by a measure of their total retail square feet of selling area. Once a calibrated production constrained spatial interaction model has been formulated for a specific set of destinations, the estimated table of such flows provides an idea of the likely inflow to each of the unconstrained destination trip ends:

\[ D_j = \sum_i P_{ij} = \sum_i A_i O_i W_j \exp(bC_{ij}). \]

The production constrained model leaves the amount and type of flow arriving at each center or store open to calculation. With such calculated inflows, the analyst has an access to a predictive model for the likely composition and size of any centers for its capture area. Think of a column of the spatial interaction matrix that leads to a specific destination as a listing of the contributions to that particular destination. Of all the flows that arrive at the destination, we may estimate the percentage that comes from each one of the surrounding regional sources. From all of those, the core or primary contributors may be determined by sorting the origins from largest to smallest and cumulating their contributions until arriving at a subset that contributes a very significant fraction of the total business of the store of interest. This is none other than Applebaum’s (1966) concept of primary trade area being the region from which a particular store draws a high percentage (say 75%) of its business.

22.2. ANALYSIS WITH RETAIL TRADE AREA MODELS

22.2.1. Spatial interaction

Spatial interaction models in general assume that interaction is determined by the attraction of the alternative facilities and by the distance separating the consumer from those alternatives. Huff (1962, 1963, 1964) and Lakshmanan and Hansen (1965) are credited with developing specialized ‘retail’ variants of the spatial interaction based allocation model. From an operational perspective, Huff introduced a practical approach to defining the ‘attraction’ of a center as the amount of floor space, rather than the population of the surrounding area as was commonly used in previous models. This opened up the interpretation of attractiveness and allowed it not only to be determined by a number of variables (e.g., number of functions, parking capacity, etc.) but also allowed attractiveness to be treated as an independent variable that could be estimated in its own right. Another major operational consideration was that Huff fitted the exponent for distance in trip-making behavior (the influence that distance has on a consumer’s store choice) to particular circumstances. Finally, he introduced a balancing term that constrained the sum of individual or zonal travel or sales to fit within an overall travel or sales limit.

With respect to the attractiveness or drawing power of a facility, Huff’s use of retail floor space has been widely adopted and adapted to include other important characteristics. Most important, though, this model demystified the idea of drawing power or attraction and allowed its direct estimation by focusing on the weight associated with it. Nakanishi and Cooper (1974) were particularly effective at utilizing Huff’s probabilistic choice framework and operational perspective to develop a linearization procedure for direct estimates of attractiveness. The MCI model is one of the best tools available for the allocation of consumer demand to facilities. The main advantages of this model is that it can incorporate a variety of attributes of the facilities under consideration by the consumer, yet it is easy to estimate. In cases where more data on the influence of various store attributes are available, the MCI model is apt to provide a more accurate estimation of market share than the original Huff model.
With spatial interaction models, then, facilities no longer have a well-defined geographic market area. Instead each store’s market area is a probabilistic surface that shows the probability of a customer from each small geographic area patronizing that facility. The exact nature of this probability surface depends on the parameters of the spatial interaction model. Incorporating spatial interaction models into a location–allocation model represents the state of the art in modeling retail site selection.

### 22.2.2. Primary trade area

Imagine a store attracting customers from surrounding census tracts or city blocks. Such data have long been analyzed by proponents of the applied school of retail trade area analysis (Applebaum, 1966). As a starting point, examine the distribution of the customers of a particular store, with regard to their origins. If the store has a weekly volume of $V$, then the customer distribution is used to spread around that demand to the originating areas, in proportion to their draw of customers. That spatially distributed demand in turn can be compared to the potential pot of money that exists in those zones available to be spent somewhere, in order to compute a measure of store penetration of the market. From the data, the top 75% (say) of the sales area may be devised, followed by the next 20% and the rest (all these are hypothetical numbers). Unless some added spatial constraints are added, it is important to note that it is not essential for the top contributing area to a store to be compact (having for example disconnected outliers).

Analytically, the primary trade area, $P$, is defined such that $\sum_{i \in P} P_{i,j} = 0.75$ and the secondary trade area, $S$, is defined such that $\sum_{i \in S} P_{i,j} = 0.20$. The remaining or ‘tertiary’ trade area, captures the remainder of the customers, often sparsely dispersed over a very wide area. For most practical purposes in the convenience sector, ‘tertiary areas’ are irrelevant to routine operations. On the other hand, significant shopping centers drawing from a large region may well have to treat the marginal sales to the edge of their tertiary area as significant ‘icing’ on the sales forecast, and may in fact be the key to understanding top-performing locations.

Retail executives are especially interested in market share, strength versus direct competitors, and in the yield of customers from a pool of potential sales dollars. It seems that the only thing worse than a store that has a small sales level is one with a large volume but under-performing its projected potential! These analyses are directed to the question: how well are our stores capturing the market? Are we leaving potential sales untapped? Or are our competitors out-maneuvering us? Penetration of the market area hinges on an assessment of how much demand is available there, and how much our particular branch is capturing.

### 22.2.3. Characterization of the demography of the trade area

The attributes and weights of demand from the particular types of respondents in the trade area can then be recovered. Say, for example, that the numbers of household in the various tracts that have particular levels of household income are given. Many useful statistics can be computed from these data. Among these are the expected values of customer characteristics over the primary, secondary, and tertiary trade areas respectively. For example, if we have a defined area that encloses the primary trade area, and the total volume of expenditure in that area is $X$, then the total volume attracted to the store of interest from within that same area is $Z$, the ratio of $X$ to $Z$ is very useful information about penetration of the market. These analyses provide the tools to diagnose practical issues in the trade area’s effectiveness, for example, by indicating untapped sales potential, the need for more intense marketing, or special circumstance arising from unique factors (ethnicity, mobility, etc.).
22.2.4. Connecting retail location models and competing destinations

Retail locational analysis is frequently carried out with the aid of spatial interaction modeling. Many features of the trade area are derived from calculations based on either actual customer origins (from a survey) or from a model of such a distribution that has been fitted from observations. In either case assume that the probability that a customer in area \( i \) shops in store \( j \) is given by \( P_{ij} \). This joint probability can be further manipulated to give \( P_{i|j} \) and \( P_{j|i} \), respectively these are:

\[
P_{i|j} = \frac{P_{ij}}{\sum_j P_{ij}} \text{ is the conditional probability that a customer who shops in } j \text{ originates from } i,
\]

and:

\[
P_{j|i} = \frac{P_{ij}}{\sum_j P_{ij}} \text{ is the conditional probability that a customer from origin } i \text{ shops in zone } j.
\]

It is this later probability that is highly useful as it allows a prediction from a given zone \( i \), of how much traffic or business might be expected to arrive at a destination in zone \( j \), and this of course can be applied either to pre-existing stores (to check model fit and validity) as well as the use of the model to forecast the likely patronage of a new or proposed location at \( j \). In that these probabilities are analytically derived from data that are exogenously available (travel times, demand expenditure parameters, and so on) they are quite easily manipulated to give forecasts of ‘what if’ for cases where there are expected changes in the data or the parameters. This kind of sensitivity analysis can provide a useful cross check on the validity of the model – for example, a sensitivity analysis should predict changes than make sense. Further, extreme values of the parameters often provide consistency checks in that the model collapses to other easily recognized forms in these special circumstances: thus a model with a distance decay parameter collapses to an all-or-nothing nearest center allocation model in the case that the beta parameter is driven to the extreme value. In this case the trade area should take on characteristics such as that seen in the ‘Voronoi’ diagram or Thiessen polygons.

In macro spatial analysis (e.g., at the scale of interregional interactions) the peripheral areas have, by definition, lower access to the dense cluster of the urban core. So, for a resident of the periphery the number of competitive alternatives in short range is comparatively small, and according to the theory of competing destinations (Fotheringham, 1983), the demand is therefore spread over few alternatives (hence is not divided up so thinly). It would be expected therefore that interaction levels over short distances are enhanced (and comparably the interaction over the longer distances is spread thinly, and hence the slope of the flow vs. distance curve is steeper than it would be expected to be, absent a spatial structure effect). At macro scales then the large beta for peripheral zones results from mis-specification, and does not correctly imply that there are larger distance decay impacts for peripheral residents; in fact, once the mis-specification is corrected, the expectation might be that peripheral residents might show a willingness to travel to distant alternatives at a rate that exceeds those of the comparatively well served central residents.

This notion of a process at one density regime being adapted for other situations was nicely foreshadowed in Berry’s (1967) classic work on commercial centers when the expected sales territory size was contrasted in low density rural Iowa with the more commercially dense built up areas of Chicago. Thus there is some interest
in whether this theory might be adapted to a more dense urban retail scenario. In the retail scenario the central or core resident has lots of alternatives within short range, and these can provide opportunity for multipurpose trips and shopping on a scale that combines multiple activities. As Eaton and Lipsey have shown, such retail agglomerations then gain more from their collocation than they lose from the presence of intensified competition. Thus the theory of competing destinations developed at a primarily interurban scale might be refined for the case of flows within an urban area, and indeed the opportunity to make multipurpose trips to clusters of shops in a city might lead to an expected agglomeration effect: what we might coin the ‘cooperative destinations’ effect arising from spillovers in retail demand (see early theory of Eaton and Lipsey, 1982).

22.3. CALCULATIONS

22.3.1. Data issues

An interesting aspect of retail trade area analysis is that the most commonly collected data (choice-based samples) are not especially well suited to direct manipulation in calibration (see a series of papers on choice based samples by O’Kelly (1999) and Ding and O’Kelly (2007)). Choice based data from frequent shopper cards at the point of sale or from check based data can tell us the distribution of actual demand around a current store. Clearly the interest in these data from a predictive point of view is to be able to use them to devise some origin based parameters such that the trade area attributes that determine the store success/failure can be studied and translated into parameters that can predict how a proposed new location (assuming that represented stores provide a decent analog for the new operation) might be expected to perform. One could expect to take data about existing operations, and develop a list of those parameters of the trade area that are expected to correlate heavily with good retail performance. The interaction model is simply an improved way to gather data and summarize standardized aspects of these trade areas to provide data about the branches. In applications, these data can then be entered into regression or other models to determine the different aspects of the trade areas that are especially highly correlated with successful operations.

An important step in managing a retail trade area data set is to understand the scope and reach of the center to the areas surrounding the store. In fundamental economic geography we learn concept of the range of the good: this is the maximum distance a customer would be willing to travel to reach the store. This maximum radius or reach has relevance for the concept of spatial interaction and trade areas as there is clearly no necessity to include demand from a place that is so far from the store as to be unable to reach that store’s trade area. Distance impedance and maximum travel radius are critical to the accurate specification of gravity models. In the case of a maximum travel radius, one has to be sure to set up a spare or ‘dummy’ destination to allow for demand that has no feasible option within range to be ‘parked’ there pending either some additional site, or some relaxation of the maximum range.

Very large energy costs cause a contraction in peoples’ willingness to travel long distance or make excess discretionary trips; instead one would expect two countervailing forces: to make a smaller number of multipurpose trips to major agglomerations would serve to support the development of a small number of heavily clustered mega malls; on the other hand the smaller willingness to travel might cause a stronger tendency to use the closer alternatives and activate the incentive to build a series of small decentralized regional centers. This trade-off between agglomeration and convenience is an interesting empirical question.
22.3.2. Determination of market effectiveness and penetration

The idea in retail interaction modeling is to use a probabilistic estimate of the demand originating in each sub-area, and its likelihood of being spent at a particular store of interest. It is convenient, though perhaps increasingly less realistic, to assume that the pool of available money is all allocated to ‘bricks and mortar’ stores, and that the demand is a simple function of the population, its income, and expenditure habits. With that assumption it is possible to take readily available census expenditure data and predict how much would be available for particular product categories in each micro-demographic area. Such micro marketing data have been used with great precision by the package goods industry, car industry, banks, and retailers in general. These applications represent one of the most powerful uses of the gravity model. Some industry specific intelligence is needed with regard to the reasonable range of potential destinations from the point of view of an origin. This is because it is necessary to be able to make an all-inclusive list of the probabilistic choice sets that exist or that might provide opportunities for the shoppers to make choices. To adapt this base case to the more realistic case of alternative non-spatial alternatives (in competition with conventional alternatives), we need to be able to estimate leakage from an origin area to electronic, catalog, and on line purchases. From the retailer’s point of view at a specific location, it is necessary to be able to circumscribe the potential originating zones from which the trip makers might be attracted. For a convenience-oriented store like a supermarket, one can imagine a reasonably compact service area. For department stores, or retailers co-located with attractions that can draw from farther places (think of Mall of America as a destination), it is perhaps a little more difficult to know the universe of the attraction, and hence difficult to make computations of the share of the attraction provided for by local or further away origins.

22.3.3. Performance assessment of existing stores

It is reasonable to assume that the primary trade area, which accounts for say 70% of the branch business is key to characterizing the stores potential customers. In an applied context, working for a retailer, we would need them to provide us with some measure for each store of the total retail volume and perhaps some breakdown by product line or class, and also an indication from the stores perspective if the chain regards the branch as successful. With the sales data we can produce measures in the surrounding zip codes for sales/household and this could give some indication of penetration rate. From that we can characterize the trade area make up for the store (Hispanic, middle class, etc.). While these data are a very big part of the puzzle, what we cannot do with such data alone is to talk about the residents in a particular subareas and their probability of being a customer. For those who are customers (AND for those who are NOT) we need some additional way to measure reasons as to why or why not. To get at these added questions we either need prior theoretical expectations, or to employ a survey to ask residents in a residential area about their reasons for shopping or not shopping at our chain. As surveys tend to be very expensive, a controlled theoretical choice experiment is perhaps a worthwhile future framework for such destination choice problems (see Eagle, 1984).

From these two sources of data detailed intelligence about the trade areas of the various branches can be accumulated and the results used to characterize the stores; if there are added data from the retailer about which stores are under- or over-performing, we could do some correlation analysis, or perhaps data envelopment analysis (Donthu and Yoo, 1998) which allows a gauge of performance vis-à-vis peer benchmarks.
22.3.4. Impact assessment

One of the most frequently asked questions from an applied perspective is to determine the loss of sales at existing stores to new entrants or competitive analysis for the diversion of existing dollars to the store of interest either from one's own chain (cannibalism) or preferably from competition.

Impacts of changed conditions are quite well accommodated by the gravity model, because the difference between two scenarios may be quite instructive. The impact of new store \( k \) on existing store \( j \), from the point of view of zone \( i \), is measured as:

\[
I_{i,j,k} = \left( \frac{P'_{ik}}{\sum \text{over all new sites } P'_{ik}} \right) (P_{ij} - P'_{ij})
\]

where \( I_{i,j,k} \) is the impact of new store \( k \) on existing store sales in zone \( i \), \( P'_{ik} \) is the new allocation to center \( k \) from zone \( i \), and \( P_{ij} \) is the allocation to center \( j \) from zone \( i \).

The types of scenario that can be handled using the methodology are as follows:

- analyze the trade areas of current stores (run with just fixed locations)
- pick sites from candidates (run with fixed and potential locations)
- re-consider current sites (make currently fixed sites flexible or optional)
- examine specific proposed sites (lock in particular new sites)
- analyze specific closings (lock out particular site and see what happens)
- analyze the opening of a known competitor (add fixed locations).

All of these versions of the problem have been deployed in practice with good empirical and quantitative results.

22.3.5. Temporal and seasonal variations in trade areas

Clearly, the volume of business is not simply related to the local demand, and the seasonal adjustment for external visitors is something that would have to be taken into account in developing accurate sales forecasts. Imagine a seaside resort such as Hilton Head, South Carolina: its sales would be quite variable over the seasons, in a cycle tied to the peak tourist demand in the northern winter. One way to do this is to examine sales records and develop a set of monthly seasonal adjustments. Whatever the base level of demand, the modeler could then devise factors to scale up or down the sales for specific months.

A simple time series model, with a set of monthly or seasonal dummy variables can be used to make an empirically fitted set of correction factors. Another way that trade area models need to be corrected is for the excess in demand that often accompanies a new store opening as the novelty of that location is added to the mix of existing stores and, at least initially, there may be large incentives or advertising efforts made to attract customers. Clearly, it would be advisable to temper these initial sales figures with some kind of decay or dilution effect that would bring the stores sales into alignment at moderate levels (see Kaufmann et al., 2000). Rules of thumb abound in this area, and equilibrium sales after opening may settle down to say 60% of the initial week sales.

22.4. LOCATION ALLOCATION MODELS

22.4.1. Introduction to location allocation models

The use of the location allocation model in retail site selection has greatly advanced over the past 15 years. Examples include the use of interaction models to develop optimal site
locations for stores in a variety of different types of retailing including supermarkets, department stores, big box retailers, and retail banking.

Successful use of these models led to their commercial acceptability and widespread adaptation in retail outlet location study (see the Thompson site selection book (Buckner, 1998)). Commercial examples in Britain include the G-MAP package (see Longley and Clarke, 1996). Specialized programs in business-GIS packages now provide routine access to methods that were previously only obtainable in customized software and research publications. This diffusion of the innovation of retail trade area analysis from specialized journals such as Environment and Planning A, into many applied sectors has been a major success for analysts. These models serve as a critical underpinning of the site selection analysis that goes into many large format stores in almost every urbanized area in the U.S. and Europe. The reason that such models are widely used is that they are essential to the rapid 'pro-forma' evaluation of numerous site proposals. The models provide the kinds of rapid computations that would ordinarily have taken a great deal of manual computation; and certainly when a chain is screening as many as 10 sites for every actual chosen location, the need for rapid analysis is obvious. For example the early studies by Applebaum (1966), directly predate the computation of trade area penetration models that may now be made using spatial interaction models.

One of the goals of this chapter is to provide the analytical background to the models that are now a commercial fact of life for retail analysis. The idea that a model of retail attraction could be deployed as a model for retail site location is an extension over the simple, earliest work in central place theory, where consumers were assumed to patronize closest centers (see also Ghosh, 1986). In turn this central place approach defined a region in close proximity to the store from which it would be reasonable to expect that the demand would be assigned to that particular store. Following a large amount of study of consumer behavior indicating dispersal of choices over many alternatives beyond just the most convenient (Clark, 1968; Hanson, 1980; O’Kelly, 1981), market researchers and others devised more precise means of estimating likely consumer behavior. The deterministic ‘all-or-nothing’ allocation of demand to the nearest or most convenient branch is no longer a necessary or indeed acceptable simplifying hypothesis about spatial behavior. Instead, we now expect that consumer behavior may be examined with the same tools that econometricians have devised for the analysis of discrete choice. Databases in turn provide a wealth of data. Geographers have derived a representation of consumer behavior with a model that locates services; this involves a breakthrough in the use of spatial interaction models. The key idea was to use the nearest center assignment of customers in central place theory, with a more realistic gravitationally based estimate of likely destination choice (O’Kelly, 1987). Thus, the customer might have a certain probability of visiting a large center that is a bit further away than a small center close to the consumer. In gauging these trade-offs, the model makes a carefully calibrated estimate of the impact of size and distance on the consumer’s willingness to travel to particular destinations. Once this calibrated model is available to us, the analyst can propose specific new site locations and gauge the expected level of consumer patronage at those sites. So called ‘turnover’ or retail sales volume is a critical first step in the analysis of any commercial property deal as the sales levels help to support the go/no go decision on rental, lease, re-model, or closing.

Location–allocation models generally involve the simultaneous selection of locations and the assignment of demand to those locations in order to optimize some specified objective or goal (usually to maximize market share or profit; see, for example, Craig, et al., 1984). These models have several advantages. They can
determine the optimal (or near optimal) location of several stores simultaneously by systematically analyzing the system-wide interactions among all stores in the market area. They are capable of utilizing a wide range of objectives that could be used in siting stores. In addition, the models are flexible in that they can incorporate the behavior of retailers, consumers and/or the retailing environment. Finally, heuristics are available for these models which provide good (optimal or near optimal) solutions and yet are easy to implement. The use of location–allocation models typically involves empirical research to determine the important store attributes for the population within the market area and a mathematical model to determine the optimal locations for retail outlets based on the pattern of market demand, store chains and existing competing outlets.8

Even though it is recognized that many consumers engage in multi-purpose, multi-stop shopping, models of multi-purpose shopping behavior have not been thoroughly integrated into facility location analysis, though early efforts by O’Kelly (1981, 1983a, b) have been recently reconsidered as the basis for new location models (Leszczyc et al., 2004). So the assumption of single-purpose trips is made in order to devise practical (usable) store-location models. Nevertheless, the fact that our analysis is primarily designed around shopping center destinations ensures that the attraction of a destination for a specific store is partly determined by the attraction of the cluster of stores as a whole.

There are several types of retail location models in the literature. Some representative examples include models which combine location–allocation with spatial interaction (for example, the MULTILOC model by Achabal et al., 1982); models which can deal with multiple objectives (for example, Min, 1987); models that consider the uncertainty inherent in the retailing environment (such as the scenario planning model by Ghosh and McLaﬀerty, 1982); and models which involve the decision maker in the decision-making process (for example, the STORELOC model by Durvasula et al., 1992). No one model is capable of handling all the important aspects of retail site selection which must be addressed in order to provide the decision maker with the best set of locations for any particular market area in which the stores will be located. Some aspects of these models are developed in more detail in the following section.

### 22.4.2. Retail location models and spatial interaction

MULTILOC (Achabal et al., 1982) was one of the first location–allocation models to simultaneously locate more than one store. The model optimizes the location of stores using the knowledge that consumers will choose among the alternatives according to a probabilistic interaction model (the MCI model). Such models maximize total profit for a retail chain (or a single store) after subtracting the fixed costs of establishing a store at the determined location (i.e., location-speciﬁc ﬁxed costs). It has later been given a more mathematical treatment in O’Kelly (1987).

The major problem facing the manager of site selection is the large number of options from which to choose, although the conceptual bases for this model are very simple. A set of potential locations is deﬁned and from this set P facilities are to be chosen. The so-called N choose P problem clearly involves a large number of combinatorial options. Not all of these choices need to be examined, however, in order for the model to make a reasonable estimate of the ideal subset of P facilities. Two major strategies are available. First, if the model can be posed as an optimization task, computer programs using mathematical techniques such as mixed integer programming (MIP) or Lagrangian relaxation to select optimal locations (O’Kelly, 1987). Second, and in many ways more robustly, the modeler can set up the problem and employ heuristics in
order to make a quick and reliable estimate of the core portion of the preferred site selections.

An example may help to make this concept clear. Suppose a clothing retailer is considering siting stores in some of the many available shopping centers in a large metropolitan region such as Atlanta. It is unlikely that the retailer would want to place a store in every available shopping center. Budget constraints would limit this option and simple common sense would indicate that the market could not bear the saturation coverage of ‘too many’ stores. The question of the optimal number of stores will be addressed presently, for now assume that the retailer has a limited number of sites that are under consideration. Therefore the retailer seeks to prioritize a subset of all the available centers that might be expected to perform well given their products and customer profile. This latter point is a key one. In order for the retailer to prioritize the store locations, the retailer needs to use an accurate model of the underlying demand for the service. Thus many geo-demographic case studies use profiles of existing customers to create a measure that reflects the attraction of the store for particular populations. This in essence is a computerized version of the classic idea by Applebaum (1966) of using analogs to project the trade area success of a proposed new store location. If the chain already has a set of stores in a wide variety of different spatial contexts, cross-sectional comparison of the performance of those stores can be used to produce a regression type model for store sales levels. Once these models are estimated, the retailer can then seek new locations where the mix of factors leans heavily towards those variables that have proven to be successful predictors in other locations. The operational version of this idea is to test each of the locational scenarios by projecting the probable trade area of each store, existing or proposed, in the context of the surrounding demographics and competition. These models have become very sophisticated because of the availability of detailed micro demographic profiles of spatial areas that may be assigned to each potential location.

As the model explores the number of locations, the analyst can keep track of the performance of those proposed sites. For example a set of five stores distributed throughout the metropolitan region might very well succeed in capturing the selected demographic submarkets that are sought and desired by this retailer. In contrast, some other combination of five stores could easily be eliminated from consideration because the sites do not deliver the expected mix and density of demand to make this package feasible. A great deal depends on a reasonable and accurate projection of the impact of each new store and its performance both against existing competitors and any stores that the chain might already have located in the district.

### 22.4.3. Combinatoric issues

A key to the efficient implementation of interaction based location models is a data structure that enables the computerized evaluation of sites to be made relatively quickly. The following notes provide a guide to the collection and organization of data in such a way as to make such computations feasible for quite a large study program. Assume that there are $M$ origin zones. The $N$ locations from which the model will select sites are organized as the columns of the interaction table with an extra column that will be used to store any user demand that is under-served by the solution program. This modification is essential when dealing with site selection models. To see this, imagine that a retailer is planning to site three new outlets in a very large metropolitan area. If the maximum distance a customer would be willing to travel to the store is set at say 10 miles (equivalent to the concept of the range in CPT) then in a large city, it is quite clear that some consumers will be too far from any of the chosen sites to be able to use this retailer’s service. It is important that
the model provide a means to calculate such unserved customers and we propose to do this by placing those ‘unserved’ consumers in a separate ‘dummy’ destination category as a holding bin for the under-provided origin zones. In the absence of competition, the goal would be to minimize unserved demand. In the presence of competitive alternatives, the goal would be to capture as much unserved demand as possible for the client’s chain.

With the exception of the concept of an additional destination, the basic calculation process is identical to that of a production constrained spatial interaction model. The device used to operationalize a particular choice of actively considered facilities is to simply keep a list of certain columns from the interaction matrix to which consumers might be allocated during that particular iteration. As the model proceeds from one locational pattern to another the set of active columns is simply switched on and off to provide an indication of the currently available destination choices. To make these calculations efficiently the computer is provided with lists pointing to various types of columns in the matrix. For example any sites which are required to be provided in all cases may be indicated by placing their column numbers in a vector of open facilities. Such a vector might be the noted by the letter $R$ for required centers. A second set of pointers might be used to indicate that in a particular analysis some potential facility locations are to be ignored completely. These, for example might be sites which we wish to lock out of the current set of optionally available sites. Yet another list could maintain a set of pointers to the available remaining unexplored options that are freely available to the model to be chosen or not as the analysis progresses. Once again having an example may help to fix these ideas. Suppose that a city currently has a total of 35 supermarkets from a number of major chain stores. One of these chains is considering a variety of expansion programs in this city. Among the locational options available to it are the acquisition of new sites, the acquisition of existing sites from competitors, and the expansion of some or all of the current stores in its portfolio. In this case it is reasonable to think that the existing stores in the market are in a sense locked in and will occur in all of the comparison scenarios: 35 columns of the interaction matrix are therefore locked in for the purposes of this initial run. Any additional locations are simply tacked on as say the 36th, 37th or 38th columns of this interaction matrix. Depending on how many candidates sites are available from which to pick these three additional locations one can imagine that the model is exploring a finite list of potential new store packages. Common sense dictates that the store chain is unlikely to want all of its new site picks in the same area, as it would make a great deal more sense to spread the chosen sites across a variety of sectors of the city. If it so happened that a pool of presently underserved demand could be found, the model would place a facility in that area. More likely, the model would be making a complex set of trade-offs, trying to eke out a market share from among and between the existing set of competitive centers, and indeed avoiding cannibalizing the existing store already owned by the chain. In this regard the strategy is essentially similar to the well known ‘gap in the map’ rubric for locating new services. The bulk of the program then would spend time computing the benefits of specific chosen alternatives in, for example, the north, east, and south suburbs. For those with the obvious question of ‘how is this done,’ it would be realistic to state that the current practice involves a combination of GIS software to manage the spatial data, customized optimization algorithms coded as executable computer programs, and a report writer to digest the output from the optimization run. While these capabilities may be combined in various customized software environments by consultants, there is probably no prepackaged comprehensive optimization environment for the applied tasks enumerated here, though this situation will undoubtedly change.
22.4.4. Heuristics and other shortcuts

The position of the store relative to the pool of demand and to other complementary and competitive stores is critical in measuring market area and size. If the objective is related to maximizing aggregate market share for our entire chain, and if there is an accurate representation of maximum distance (reservation distance) we can expect that the model will ‘naturally’ space out our stores giving them somewhat non-overlapping exclusive market areas. Nevertheless, when two stores are close enough to contest a middle ground then the gravity model will do better than the usual deterministic all-or-nothing location models. The gravity model will in fact partition the demand between the centers in proportion to their attraction and weight.

If such a model is to be run in site selection mode, realize that the attraction/repulsion score will have to be computed for prospective as well existing sites – in other words it has to be some calculated feature of sites that are ‘prospects’: it cannot be simply some observable feature of existing sites.

Required site
These are locations of our own chain that we wish to keep. We can also, in some cases, represent the locations of competitor stores that we know are remaining in operation. These would be treated as fixed sites.

Prohibited site
Areas or store locations that are prohibited from entering the model are equally important to a realistic implementation. If we are sure that the chain does not wish to enter certain malls, or if the location in proximity to some existing stores is strongly discouraged, then candidate locations in the ‘no go zone’ should be flagged to (a) save computer time; and (b) and to enhance the chance that the model will focus attention in areas that are worth investigating.

Flexible sites
The set of locations from which the model will pick are predefined by the user. These could be the result of selection set operations, query based lists, or geographically delimited regions on the screen. What is important is for an underlying comprehensive data base to be kept up to date in order for the analysts to have meaningful choices from which to derive the set of active alternatives.

22.4.5. Computable Location Models

Location model must be flexible to allow analysis of different scenarios. The model takes as input the required and flexible sites. The existing literature contains several models dealing with joint location and allocation under spatial interaction: these however, need to be modified to handle realistic selection sets of required and prohibited sites.

The best practice at this time is to use a robust vertex substitution method appropriately modified to handle lists of required and prohibited sites, as well as efficiently managing the introduction of new candidate locations.

The vertex substitution method also needs to include the capability of a maximum service radius for the facilities, and for this radius to be flexible/variable between centers: this is essential if some notion of center hierarchy is to be accommodated. It should be clarified that the vertex substitution method is a local optimally solution in the sense that there may be a better solution that was not reached during the course of the exploration; this possibility can be reduced by trying the method with various starting values. Research experience has shown, however, that the good locations ‘stand out’ very well and the possibility that the vertex substitution method completely misses the best package of locations is remote. One idea that is suggested to prevent mistakes due to local optimality is to produce a list not only of the best locations but other close contenders
discovered in the course of the algorithm’s progress.

Research by Church has shown that the introduction of maximum service radii into a median type of problem (which is what we have) disrupts one ‘normal’ property of the model, making it potentially possible that the ‘optimal’ locations occur at points other than the nodes of the network. However, the actual problem that we are concerned with realistically limits the feasible locations to the nodes of the network, as this is where the shopping centers are. In other words we ignore the theoretical possibility that the true optimal solution is at an intermediate location along street segments, as in practice this kind of locational solution would not be permissible.

What does experience tell us about the solution of location allocation models? The basic model is conceptually very simple and easy to understand. The idea is to systematically explore alternative locational scenarios. The method takes as input the fixed locations, the candidates, and the prohibited sites (if any). As output the model produces the requested number of additional facility sites, and reports on the area characteristics of both the current and the new sites. The candidates are either a comprehensive list of all feasible shopping centers, stores are generated from a list of ‘picks’ and potential sites. The user may select the candidates as those sites which meet some criteria, and the detail and realism of these selection criteria are really only constrained by the imagination of the user. All kinds of filters can be used, including center size, or selections can be based on attributes of the centers. Having selected the candidates, the user would have to select the objective function: normally this is driven on the basis of aggregate market share, or demand, or minimizing competitors share. This is potentially extended to include acquisition, lease, closing and opening financial decisions.

Vertex substitution has the great advantage that as a general purpose optimization strategy (i.e., heuristic) it is robust to changes of objective function, in a way, for example, that would not be true of a specialized exact optimization code. In other words, the weakness of an ‘exact’ method is that it typically has to exploit some aspect of the problem structure and any change in that structure would likely undermine the mathematical formulation. Heuristics (and there are many of these available for combinatorial problems) can frequently be set up to explore a solution space effectively and this can be chosen to evaluate the users choice of objective (and indeed multiple objectives) to achieve the desired goals. Indeed the final great advantage of an exploratory heuristic is that by careful book-keeping many ‘runner up’ or close alternative solutions can be kept and compared.

22.5. STRATEGIC PLANNING EXAMPLES

22.5.1. Shopping centers

Store location siting is often made from among a predefined set of existing shopping centers, so in a sense the set from which the strategic location is to be chosen is already fixed. Thus, 1747 block groups in Atlanta represent the pools of available demand, which for the purposes of this simple example are weighted by the population or disposable income as a proxy for the demand. Assume a chain has 12 existing stores distributed throughout the Atlanta region in specific shopping centers. There are approximately 230 potential sites in shopping centers. Assume that the reach or ‘draw’ of the center candidates is a function of the size of the center – in other words the decision to open a new branch in a thriving center with a super-regional draw might be appropriately measured by using the size of the center as a proxy for its suitability. Suppose then that the location allocation model algorithm, such as the Interchange Heuristic, picks four locations as the close-to-optimal added sites. (We are careful not to call them optimal in view of the many simplifications and the use
of a heuristic which after all depends on some short cuts to avoid complete enumeration of the many thousands of combinations that are available.

The impact of each new center on the 12 existing sites is then operationally measured using a formula such as the one discussed above.

### 22.5.2. Chain combinations

Sales of branches in two existing sets of chain stores can give a good clue as to the best ones to keep in the combined operation, but that still leaves a difficult problem to determine which ones to close. Predicting retention of customers from old stores to re-aligned new branches is also difficult though the managers of such operations may have good insight into the likely levels of customer loyalty.

An interesting question is to determine the diversion of sales or the result of a store/chain closure. Such questions frequently are presented in practice to retailers as they have the option to purchase competitors sites. Which of these sites would make good acquisitions (if the option to cherry pick the best of the available store)? Which would be blended well and open under the new label if the acquiring chain gets the whole suite?

If two chains merge, and there are regulatory concerns that the two chains have to divest some of their branches, or wish to streamline their combined operations, one would have to analyze the closure of branches one by one to determine the package that makes the most sense from the point of view of the combined operations.

### 22.6. SUMMARY AND CONCLUSIONS

The great strength of the gravity model is its simplicity and its allocation of demand to centers in proportion to their attraction and inversely proportional to distance. It can incorporate center specific attraction and center specific maximum trade area radii.

The strength of the SI based location model is that it provides assistance with all of the

<table>
<thead>
<tr>
<th>Where did the impact on store come from?</th>
<th>NEW1</th>
<th>NEW2</th>
<th>NEW3</th>
<th>NEW4</th>
<th>Taken from</th>
<th>store number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.04</td>
<td>1</td>
<td>0.04</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1.02</td>
<td>0</td>
<td>0.47</td>
<td>2</td>
<td>0.47</td>
</tr>
<tr>
<td>0</td>
<td>0.65</td>
<td>0</td>
<td>0.01</td>
<td>0.41</td>
<td>3</td>
<td>1.03</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>1.06</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1.11</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0.04</td>
</tr>
<tr>
<td>0</td>
<td>0.22</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0.02</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>9</td>
<td>1.11</td>
</tr>
<tr>
<td>0.28</td>
<td>1.67</td>
<td>1.62</td>
<td>0.93</td>
<td>From existing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.82</td>
<td>2.35</td>
<td>2.64</td>
<td>2.78</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.54</td>
<td>0.68</td>
<td>1.02</td>
<td>1.85</td>
<td>Net added</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>Total share</td>
<td>0.9434</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After</td>
<td>Total share</td>
<td>0.9842</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
following tasks: measuring saturation, impact of changes on current trade areas, assessment of the advantages of certain locations for particular formats, and an estimation of the forecast of sales. In addition the allocation models allow a profile of the demographics of a trade area.

What would take a large amount of extra research effort, but which in my opinion would be well worth while, would be the inclusion of the interaction based model in a multiobjective and multiattribute decision framework. The difficulty would be to elicit from the decision maker a set of trade off parameters that define the relative scales for the attributes of the alternative locational packages.

The mechanism reviewed in this chapter will operate to allocate the sales from the origin zones to the destinations is called the allocation model. It is driven by a gravity based spatial interaction model, and given careful data and careful assessment of the foundation assumptions this is a robust model for trade area delimitation.

AKNOWLEDGMENTS

Parts of this chapter are based on materials developed over many years in my Retail Location Seminar where comments from Debbie Bryan, Tony Grubesic, and Tim Matisziw are gratefully acknowledged (see specific footnotes). In addition, a great deal of the common sense application flavor of this paper derives from conversations with Jim Stone (GeoVue), Tony Lea (Environics Analytics), and Steve Wheelock. I thank these individuals while taking full responsibility for the product here. Some material originally prepared as a discussion/research memo on location models for Geonomics. See 1.2, 4.5, 4.6, and examples in 5. Other material derived from ‘Retail Location Models and Spatial Interaction’ M.E. O’Kelly and D. Bryan. A Review of Modeling in Retail Location Unpublished working paper.

NOTES

1 Introduction is based on Geography 845 Lecture Jan 2, 2001.
2 A major sector using the results from spatial modeling capability is that of businesses with multi-store/branch locations. Home Depot for example has made extensive use of reports from what used to be Thompson Associates, and is now a unit of MapInfo in Ann Arbor Mt. Other well known users include McDonalds and Blockbuster.
3 Based on applications as discussed with Jim Stone and Tony Lea.
4 Some aspects of these following paragraphs have benefited from discussion with Jim Stone.
5 The target level of goodness-of-fit in convenience store forecasting models is for high r-square values (about 0.8).
6 Section 22.2.1 is based on ‘Retail Location Models and Spatial Interaction’ by M.E. O’Kelly and D. Bryan, A Review of Modeling in Retail Location. Unpublished working paper.
7 GeoVue has a gravity based software package. ESRI Business Analyst software has a Huff trade area model.
8 This material derived from ‘Retail Location Models and Spatial Interaction’ by M.E. O’Kelly and D. Bryan, A Review of Modeling in Retail Location. Unpublished working paper.
9 Material in section 4.5 was originally discussed in an explanatory memo from this author to Jim Stone at Geonomics (now GeoVue). Jim’s critique was helpful in framing the discussion.
10 This section benefited from discussion with Jim Stone and Steve Wheelock.

REFERENCES

(Although not all these papers are cited directly, these are however all influential papers in my analysis; they are retained as a general bibliographic resource.)


