2003-01

Assessment of business potential at retail sites: empirical findings from a US supermarket chain

Smith, L. Douglas; Sanchez, Susan M.

Routledge

Smith, L. Douglas, and Susan Sanchez. "Assessment of business potential at retail sites: empirical findings from a US supermarket chain." The International Review of Retail, Distribution and Consumer Research 13, no. 1 (2003): 37-58.
http://hdl.handle.net/10945/59243

Int. Rev. of Retail, Distribution and Consumer Research 13:1 January 2003 37–58

Assessment of business potential at retail sites: empirical findings from a US supermarket chain

L. Douglas Smith and Susan M. Sanchez

Abstract

We address the problem of estimating retail business potential at alternative sites, with concern for assessing performance relative to potential in existing markets and for identifying the best sites for expansion into new markets. At question is the utility of information typically used in formal retail patronage models, in comparison with additional information considered important by retail executives. Relevant data are gathered from secondary sources and intensive in-store surveys are conducted to produce a portfolio of information about neighbourhood demographics, store ambience, variety and quality of products and services, relative prices of selected products, etc. for stores in a retail grocery chain and competitive stores in the chain's markets. We experiment with alternative statistical models for store performance to determine the consequence of restricting the types of data available when constructing the models. Our findings suggest that while information about store location and surrounding areas, store characteristics and competitive position are all required to obtain the best assessment for business potential at a site, a few key variables on each dimension offer the bulk of explanatory power. Further, the spatial-locational variables affect all measures of store performance in intuitive directions, whereas the effects of other variables differ according to performance measure and reflect the store's market position.

Keywords

Retail location, store choice, competition, site selection, business potential

Introduction

The importance of physical location to a retail outlet is well recognized. The assessment of business potential at a site is, however, a challenging process. One approach to the problem involves two basic steps. First, demographic statistics and historical sales figures are used to estimate the total demand for a general product or service that will emanate from defined geographic areas. Next, the total

L. Douglas Smith, Center for Business and Industrial Studies, College of Business Administration, University of Missouri-St Louis, St Louis, MO 63121, USA; Tel: +(314) 516 6108; Fax: +(314) 516 6827; Susan M. Sanchez, Operations Research Department, Naval Postgraduate School, Monterey, CA 93943, USA.

The International Review of Retail, Distribution and Consumer Research
ISSN 0959-3969 print/ISSN 1466-4402 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0959396032000051684

business from each area is allocated among alternative retail outlets according to their proximity and other characteristics. Such allocations may occur with the help of spatial-interaction (gravity or multiplicative competitive interaction) models for predicting market share (Achabal *et al.* 1982; Bucklin 1971; Cooper and Nakanishini 1988; Ghosh and McLafferty 1987; Huff 1963; Lundsten 1978). They may also include capacity as an explicit constraint (Smith and Moses 1996). Another approach to the problem is the direct application of statistical models, which include variables that affect the total consumption of products and services by a residential or transient population, as well as variables that affect the choice of retail outlet (Agrawal and Schorling 1996; Mahajan *et al.* 1982). A third approach to facility location pertains to supply-chain considerations and often employs mathematical programming to select the best combination of facilities (Tyagi and Das 1997). The optimization models, however, require estimates of the amount of business that will accrue at alternative sites as the retail nodes in the distribution networks are changed.

Research on market share for individual products has demonstrated the influence of numerous factors in the selection of a retail outlet for a purchase. Mitchell and Kiral (1999) found that consumers perceived a tradeoff between value-for-money and attributes such as store attractiveness and the variety and quality of goods. Time and convenience also affected the choice of outlet. Accessibility, product assortment, service and price were shown by Tantiwong *et al.* (1985) to have varying influences and interactive effects for elderly shoppers. Goodwin and McElwee (1999) investigated the desirability of various service and product factors, as well as how these differ with age for mature adults. Mulhern and Williams (1994) found that ethnic groups differ in tendency to purchase store brands. Koelemeiger and Oppewal (1999) confirmed that interactions among assortment, ambience and distance to competitive stores affect purchase patterns of a specific product. Bodkin and Lord (1997) found that less cross-shopping tends to occur at centres dominated by a few large 'superstores' than in other types of centres. Guerts and Whitlark (1993) have related sales of a product to marketing variables such as product features and packaging, advertising, personal selling programmes, distribution channels and pricing. Simonson (1999) discusses ways in which product assortment can shape customer preferences.

Managers' perceptions of information relevant to retail site location

While there have been great advances in the development of formal analytical models and databases for retail site selection, Clarke *et al.* (2000) have noticed that managerial intuition continues to play a major role in the selection of sites for new retail outlets. They assert that 'existing models for retail site evaluation ignore the useful "soft" insights of retail executives'. Further, they question whether the results of analysis using formal models might often be at odds with the knowledge of experienced managers, thus accounting for the frequent reliance on subjective judgement. Considering the arguments of Blattberg and Hoch (1990), Clarke *et al.* suggest that some combination of quantitative modelling and managerial intuition should give the best assessment of business potential at a site. From an intensive 'cognitive mapping' exercise with experienced managers, they identify factors (some objective, others subjective) that represent the information used to form the

intuitive judgements of managers. They suggest that incorporating related qualitative information in the formal modelling approaches might result in better forecasts.

The cognitive mapping exercise of Clarke *et al.* (2000, op cit.) revealed that managers of do-it-yourself retail outlets consider each of the basic factors that typically appear in spatial-location models (namely characteristics of the population that affect demand for the store's products; the size of the store; and the size and proximity of competitive outlets). One primary node in the cognitive map is the 'quality of the catchment area'. The related variables (size of population, population density, characteristics of housing stock and ownership, age distributions, housing prices and employment rates) are available from governmental and commercial sources. A second primary node in the cognitive map is 'access to store'. The related variables (distances from population centres, convenience of access from major arterial roads, sign-posting, natural traffic flows in the vicinity of the store, size of parking lot, convenience of store entrance to parking areas) are much less amenable to objective quantitative measurement and therefore more reliant on aggregate subjective assessments. A third primary node is 'store characteristics'. Included here are size of selling area, store layout, ease of shopping experience, range of products, numbers and affiliations of competitors, historical presence and strength in the market, age and size of stores, and the presence of other types of retail outlets. Also mentioned were the quality of service delivery as affected by the quality of management and staff. The managers were thus found also to consider factors identified in the published research on market share and consumer choice for individual products and services. Some of these variables are objectively measurable; others require subjective assessments. From the managers' cognitive maps, Clarke *et al.* summarize the basic determinants of store performance as:

1 site location;
2 store operations;
3 competition.

They see managers as first considering the fundamental market potential afforded by the location, and then considering the ability of the store to gain market share in light of the competition in the area. The assessment of competitive position would begin with a consideration of the spatial-locational aspects and extend to characteristics of the stores, with their physical characteristics and mixes of products and services.

The 'locational modeling system', created for analysis in conjunction with Tesco's systematic programme of expansion of its chain of supermarkets in the UK, has been described by Penny and Broom (1988). The range of generic variables mentioned for siting the grocery supermarkets is quite similar to those revealed in the cognitive mapping exercise for the do-it-yourself centres. Considered in the Tesco system are variables that describe the site, store, attributes of the population in the catchment area, and indicators of competition (extent and quality). The need for on-site assessment of competitive stores (as opposed to reliance on secondary sources) is mentioned. Incorporated into the formal MCI model for estimating market share is simply a general measure of store size and quality, but the aggregate assessment of store quality (attractiveness) is still made with the help of detailed information collected from direct observation at individual stores. Fotheringham (1988), in contrast, when

estimating business potential for grocery stores in the USA, expects diminishing returns from the types of information collected in detailed store surveys and doubts the value of much of the detail derived from in-store surveys when the concern is for choosing a site. He advocates the use of 'parsimonious' models that reflect the spatial distribution of customers, spatial distribution of stores, store size and racial composition of the population in the catchment areas when evaluating the fundamental business potential at a site, and then giving consideration to more detailed information when analysing sales for individual products and developing related marketing strategies. Richards (1984) describes the variety of sources of demographic information in the USA and in the UK for estimating retail sales potential. He asserts that total population, income levels and family size are the critical variables for estimating the amounts spent at grocery supermarkets.

Based on the aforementioned studies of store performance, published research on consumers' choice of retail outlets, and interviews with senior retail executives, the key factors judged to determine success at a retail site can be summarized as:

1 population characteristics related to consumption of goods and services;
2 accessibility and sizes of competitive retail outlets;
3 comparative prices of goods and services;
4 variety of goods;
5 quality of goods;
6 shopping ambience;
7 cross-shopping opportunities.

For a chain interested in systematically expanding into new markets, assessing performance relative to potential in existing markets, or in possibly withdrawing from existing markets, consideration would be given to combinations of objective and subjective data for these factors. There is controversy, however, about the incremental value of various types of information in assessing business potential at a site. The collection of comparative data for all of these characteristics can be extremely costly and time-consuming. In addition, items 3, 4 and 5 can change quickly in response to competitive pressures. How beneficial is it to incorporate the various types of quantitative and qualitative information in formal models used to assess retail potential at a site? Can formal models, so refined, help in the development and assessment of competitive strategies? Is it possible to collect information on a few key items regarding the general characteristics of the store, its surroundings and existing competitive outlets, and to rely on that selected information to obtain good estimates of the business potential? In this paper, we explicitly address these questions with the help of data collected for a chain of discount supermarkets with outlets from coast to coast in the United States. This work was originally undertaken to address the very practical concern of identifying the best sites for expanding the chain's network of stores. The study, however, also enabled us to investigate the utility of various types of objective and subjective information, which are claimed by academicians and professional managers to be especially relevant when estimating retail business potential.

Research methodology

Our approach to the problem is to undertake an intensive compilation of information that prior studies and managerial staff identify as relevant to the

determination of the overall performance of a store. We collect the data for existing stores, construct statistical models to explain several measures of store performance, and examine how the explanatory power of the models changes when we eliminate different types of information. Standard demographic data are obtained from secondary sources. Information about stores and their surroundings are gathered from a survey of grocery shopping opportunities that was designed in collaboration with corporate managers and marketing staff and administered in 135 corporate markets in 18 states. The survey includes both objective and subjective elements for a retail grocery outlet store and its environment. Marketing executives and marketing staff helped to design the survey, ensuring that it incorporated information which they judged as most important in determining store performance. Data for the study were gathered for a total of 655 stores (135 corporate stores and 520 competitive stores). The survey first identifies the store by name, physical location and affiliation (whether part of a chain or independent). The type of outlet (grocery store, discount store, convenience store or ethnic food store) is noted along with the store's marketing position (high–low, everyday low price, limited assortment discount, limited assortment upscale, or upscale). Marketing executives for the chain chose these categories as most relevant for the business at hand. Cross-shopping opportunities in the immediate environment are identified and the aesthetics and safety of the surrounding neighbourhood are assessed by observation (with neighbourhood crime rate being described by a store employee as low, medium or high). The accessibility from traffic arteries is rated on a semantic differential, and several indicators of store capacity are measured. The variety of several products and services, apparent quality and freshness of products, and general store and service characteristics (attractiveness in presentation of merchandise, efficiency of checkout and friendliness of employees) are rated by the observers as low, medium or high. The availability of a number of services is registered as 'yes' or 'no' in each case, and the internal and external appearance of the store are rated on a semantic differential from 1 to 5. The lowest prices at which each of 15 staple products is sold in the store (at the time of the survey) are recorded (regular and sales price, where relevant).

Written guidelines for completion of the semantic differentials were provided to each interviewer. Surveys within a specific market were conducted on the same day by a single individual. This ensured consistency when comparing prices of products and when computing relative measures of attractiveness for the different stores in a market. As part of the training process and to test for inter-rater reliability when using the semantic differentials, we had more than one survey conducted on a total of 11 occasions at 8 different stores. For stores where more than one survey was conducted, the maximum difference in recorded scores for any particular item was one point. This affirmed that the surveyors were interpreting the guidelines on a consistent basis. In cases of duplicate observations for a store, we simply retained the first set recorded on the database when preparing data for the regression analysis. From 666 surveys, we thus obtained data for 655 stores in 135 markets.

Using the raw data, we compute summary statistics for each corporate market that compare the attributes of corporate stores and their locations with the attributes of competitive stores in the corresponding market area. The distance to nearest competitor, average distance to all competitors, average values for each of the attributes recorded on the survey, and the ratio of each corporate measure to the

average of the corresponding measure for all 'relevant competitors' are computed. We note whether the nearest competitor follows a high–low, everyday-low-price, or limited-assortment-discount marketing strategy and produced corresponding dummy (0–1) variables. We compute the average price of the 15 test items at each store, count the number of items for which the price was lowest in the corporate store, and compute the percentage saving that would occur if the goods were purchased at the corporate store instead of at a competitor in the market area. These variables are merged with a series of demographic statistics for populations within various radii of the corporate store's location and with a series of performance statistics for each corporate store. From the raw data, we create sets of predictive variables and categorize them in five groups as follows:

1 Demographic characteristics of the residents in the vicinity of a store location. The demographic data are readily available from governmental and commercial sources.
2 Objective characteristics of the corporate store such as selling area, hours of operation, etc. These data are available in corporate records and conveniently verifiable by regional sales managers.
3 Objective characteristics of the competitive stores in the market area and, in addition, the distances between each competitive store and the corporate store. These data required direct observation in the market areas. Although some information is available in commercially available databases, the reliability and completeness of information tended to vary considerably among market areas. Nevertheless, these types of data are able to be collected quite efficiently. Only a few minutes are required at each site. Travel time among sites is the primary expense here.
4 Store characteristics that require a subjective assessment, such as convenience of access from transportation arteries, neighbourhood ambience, and range and quality of goods and services. The collection of these data requires intensive on-site observation and assessment, with 10–20 minutes required at each store.
5 Comparative prices for a basket of staples in the product line. This extended the time on-site by approximately 15–35 minutes, as occasionally it was quite difficult to find the specific products in some stores and incidental time was often required to visit with store personnel.

Because there is considerably more expense in gathering data first-hand, we wish to produce an indication of the general explanatory power that can be derived from each type of data alone and from various combinations of data types. In succeeding sections, we describe the analytical database, show the results of creating models with different blocks of variables, and draw inferences about the utility of various types of information in estimating different measures of store performance.

As mentioned above, primary data were collected for 655 stores in 135 corporate markets in 18 states. The markets were chosen by corporate marketing executives to give a broad cross-section of the chain's own (non-franchized) stores. To help generate a list of possible competitors to each corporate store, we used the PRO CD geographic software, which contains names, addresses, telephone numbers, zip codes, and longitude and latitude for businesses listed in white-pages and yellow-pages telephone directories. Lists of grocery outlets from the PRO CD geographic information system were merged with corporate competitive reports

and presented to district managers, who confirmed or refined the list of 'relevant competitors' for each market area. In a few instances, a competitive store was considered to be a competitor to two corporate stores. To handle such cases, we used the information for that store in both markets (except distance, which was re-computed for the second market). Longitude and latitude were obtained directly from the PRO CD database, using the street address for stores listed there. For stores without local telephone listings, the longitude and latitude were obtained from street addresses by using the MAPBLAST software on the Internet. Maps were produced to facilitate the verification of physical locations of stores during administration of the surveys.

Results of analysis to determine the utility of various types of information

We constructed statistical models for six alternative measures of store performance. Defining a 'month' as a consecutive four-week period, the measures are:

1 Average of monthly sales during the 13 'months' of fiscal 1998;
2 Average of monthly sales per square foot;
3 Average sales per customer visit;
4 Market penetration (average monthly store sales as a percentage of estimated monthly expenditures on grocery products in the market area);
5 Sales growth ratio (average monthly sales during the last 7 months of fiscal 1998, divided by average monthly sales during the previous six months);
6 Average monthly profit contribution (after extraordinary items and interest expense but before taxes).

To explain the variation in these measures among the different corporate stores, we constructed regression models using the aforementioned blocks of variables that describe demographics of the market areas, store characteristics, and competitive characteristics. The following explanatory variables were selected as candidates to include in the regression models.

Demographic variables

The basic demographic variables used in our analysis were:

- Estimated total monthly expenditures on food by residents.
- Demographic variables used in estimating expenditures on different food groups:
  - Percentage of residents living below the poverty line;
  - Percentage of workforce unemployed;
  - Median family income;
  - Percentage of population over 18 that has not completed high school;
  - Median residential property value.
- Ethnicity of population served:
  - Percentage of the population that is black;
  - Percentage of the population of Hispanic origin.

Estimates of monthly food expenditures were purchased from a commercial source that based its estimates on survey information in conjunction with demographic data. We included individual demographic variables in our models, as well, to determine whether the chain's stores had more appeal to some demographic segments than others, and also to see whether the estimated food expenditures provide significant incremental information beyond the demographic variables that we selected. Ethnicity was specifically included because aforementioned buyer surveys show that choice of merchandise and retail outlets often differs among ethnic groups. Demographic data were compiled for 1-mile, 2-mile, and 3-mile radii about the store, and also for variable radii that reflected the actual distances from which customers are normally drawn to the store. The variable radii were stated to the nearest mile and were derived by the corporate marketing department considering the locations of the primary competitors and the addresses registered in surveys and contest drawings that are conducted periodically in the individual stores. For each of these demographic variables (and for 1997 population), we computed the ratio of the 3-mile to 1-mile values as an indication of the rate of change in the characteristic as the market area is expanded about the store. We refer to that set of variables as gradient variables or 'gradients'. The gradient variables were intended to measure the extent to which the characteristics of the population in the immediate vicinity of the store differed from the population in the surrounding area. We also produced an additional indicator of urbanization or market density (DENSEMKT) according to whether the total weekly food sales by residents within 1 mile of the store's location were estimated to exceed $30,000 (an amount exceeded in about 25% of corporate markets).

Store characteristics

To describe the corporate store characteristics we included the following variables from the store surveys:

- Service capacity (size);
- Total weekly hours of operation;
- Selling area (square feet);
- Number of parking spots;
- Shopping ambience;
- Neighbourhood appearance;
- Neighbourhood crime rate;
- Security in store and parking area;
- Rating of internal appearance;
- Rating of external appearance;
- Number of cross-shopping opportunities;
- Variety of merchandise;
- Quality of goods and services;
- Attractiveness of presentation of merchandise;
- Efficiency at checkout;
- Number of services offered;
- Accessibility to the corporate store;

- Visibility of store from major transportation artery;
- Ease of access from major transportation artery;
- Incidental variables proposed by management;
- Whether fresh meat is sold;
- Whether the store was new in the last 6, 12, 18, or 24 months.

Competitive characteristics

To describe the competitive environment, we used summary statistics from the surveys of the competitive stores in the market area for a corporate store and produced relative measures by taking the value of a variable for the corporate store and dividing it by the average value for competitive stores. The resulting set of variables included:

- Accessibility of competitors
  - Total number of competitors to the corporate store included in the variable-radius market area (represented variously as the number itself, the inverse of the number, and as total food expenditures divided by the number of competitors);
  - Average distance (miles) from the corporate store to the competitive stores;
  - Minimum distance (miles) from the corporate store to a competitive store;
  - Whether there is a limited assortment discount store within 1 mile;
  - Whether there is a limited assortment discount store within 3 miles;
  - Relative visibility of store from transportation artery;
  - Relative ease of access.
- Relative prices
  - Average percentage saving on the 15-item basket of staples, in comparison with the minimum price charged by any competitor in the market area (AVSAV1);
  - Average percentage saving on the 15-item basket of staples, in comparison with the average price charged by competitors in the market area (AVSAV2);
  - Number of items in the basket of staples for which the corporate price was lowest;
  - Whether the marketing position of the nearest competitor is high–low;
  - Whether the marketing position of the nearest competitor is everyday-low;
  - Whether the marketing position of the nearest competitor is limited-assortment-discount;
  - Relative number of hours of operation;
  - Relative size of selling area;
  - Relative number of parking spots.
- Comparative shopping ambience
  - Relative rating of internal appearance;
  - Relative rating of external appearance;
  - Relative level of crime in the immediate area;
  - Relative security provided to shoppers;
  - Relative variety of products.

- Relative quality of merchandise and services
  - Relative attractiveness of presentation of products;
  - Relative quality of fresh meat;
  - Relative quality of fresh produce.

Originally, the number of checkout stations was included as an explanatory variable for the store characteristics and competitive characteristics. This variable was among the strongest predictors of performance. Although the number of checkout stations is a critical determinant of store sales capacity, we removed this variable after discussions with managers who revealed that checkout stations in several corporate stores had been added in response to demand. We thus have chosen to ignore, in this study, the circular relationship between the number of checkout aisles and sales volumes.

To assess the utility of the various types of variables in estimating store performance, we created regression models that included different blocks of variables and compared their explanatory power. Recognizing that several measures may convey similar information (average income and percentage of residents with higher education, for example), we took two approaches to dealing with collinearity among the explanatory variables. In the first approach, we simply dropped a variable from the model if its marginal contribution was not statistically significant (i.e. if its associated t-test was not significant at the 0.1 level). The actual p-values for most of the 'surviving' variables were, however, much smaller, thus meeting a much more stringent test. Most variables that were retained in the models were statistically significant, on the margin, at the 0.01 (or lower) level. A representative holdout sample of 12 stores was used to test the comparable performance of the models in market areas that were not used in calibration of the models. In a second approach to the issue of collinearity, we performed factor analyses upon three groups of explanatory variables (demographic variables, store characteristics and competitive characteristics) to create, in each case, sets of orthogonal factor scores for the respective characteristics. Alternative regression models were created using the factor scores instead of the individual variables.

In the first instance, we limited the set of possible explanatory variables to the demographic variables (and gradients). Next, we restricted the choice to 3-mile demographic variables and gradients because the extent of customer draw at a new site would be unknown. The process continued, with different combinations of variables allowed. In the first section of Table 1, we summarize the percentage of variation in performance (R-squared) that was explained by the model in each case.

It is evident, from the results in Table 1, that each of the types of variables contains significant information that can help to explain or predict store performance in a market area. Demographic data alone explained about 60% of variation in sales, sales per customer, and market penetration among the corporate stores surveyed. If racial characteristics of the population were ignored, the percentage of variability in sales and sales-per-customer dropped significantly, suggesting that the chain is geared better to attract some ethnic segments of the population than others. Perhaps because the characteristics of corporate stores were quite uniform, the demographic characteristics alone explained a larger percentage of variation in performance than do the corporate store characteristics alone. Competitive characteristics seemed to have about the same explanatory

Table 1 Proportion of variance in performance measures explained by different variable categories (using 0.1 level of statistical significance for inclusion of variables in each group)

Variables included	*Sales*	*Sales per square foot*	*Sales per customer*	*Market penetration*	*Sales growth*	*Profit*
All demographics	0.58	0.51	0.59	0.70	0.28	0.50
3-mile demographics and gradients	0.53	0.46	0.52	0.48	0.28	0.42
V-mile demographics and gradients	0.51	0.43	0.52	0.62	0.27	0.39
Nonracial 3-mile demographics and gradients	0.31	0.33	0.38	0.38	0.17	0.27
Nonracial *v*-mile demographics and gradients	0.34	0.34	0.41	0.58	0.18	0.28
Store characteristics *v*-mile demographics and gradients	0.60	0.55	0.61	0.60	0.57	0.56
Store characteristics	0.44	0.42	0.35	0.09	0.38	0.38
Competitive characteristics	0.43	0.31	0.41	0.53	0.24	0.37
Store characteristics, *v*-mile demographic variables, gradient variables and competitive characteristics	0.81	0.85	0.75	0.87	0.58	0.75
Parsimonious model using all variable categories	0.65	0.66	0.64	0.71	0.54	0.63
Parsimonious models using non-survey variables	0.60	0.56	0.62	0.66	0.34	0.46
Demographic factor scores	0.34	0.29	0.46	0.43	0.12	0.24
Demographic and store factor stores	0.49	0.44	0.58	0.44	0.25	0.43
Demographic, store and competitive factor scores	0.64	0.57	0.64	0.59	0.31	0.56

power as store characteristics. All characteristics, in combination, explained over 70% of the variability in sales and profitability. (They explained over 80% when the number of accessible checkout lanes was included.) The parsimonious models involving variables from each block provided essentially the same explanatory power as the corresponding models that contained 'moderating' variables. Removal of the survey information from the parsimonious models, however, caused a more substantial decrease in the explanatory power of the models for sales growth and profit than it did for the models pertaining to sales, sales per square foot, sales per customer, or market penetration.

In the factor analysis for each group of variables, we extracted principal components and selected factors for which eigenvalues exceeded 1.0. Five demographic factors emerged which explained 72% of the total variance in the demographic variables; eight store factors emerged which explained 73% of the variance in the store characteristics; and 11 competitive factors emerged which explained 72% of the variance in the competitive measures. Employing these factors instead of the individual variables, we constructed regression models for the six performance measures and eliminated any factor scores that were not statistically significant on the margin. The last section of Table 1 contains comparisons of the explanatory power for resulting models as additional factors of each type are incorporated into the models. Again, it is evident that the chain's own store characteristics added explanatory power to models that contained demographic data alone and competitive characteristics added explanatory power to models that contained both demographic data and the chain's own store characteristics. Because about 28% of the variation in the individual variables was lost when producing the factor scores, the explanatory power of the regression models produced from the factor scores was lower than that produced from the underlying variables.

Parsimonious models of expected store performance

In the next stage of the analysis, we undertook to identify key variables that help to explain or predict the measures of store performance at a site and to assess their relative importance. We constructed another set of parsimonious models for this purpose by employing a one-tailed test at a .05 level of confidence to eliminate any variable that clearly appeared to have an impact that is counter to theory or expectations (i.e. an 'incorrect' sign). For this set of parsimonious, or predictive models, we considered variables from all categories including variable-radius demographic variables and gradients, store characteristics, and competitive characteristics. (For sales growth and profit we also included variables that indicated whether fresh meat or wide frozen displays were added at the corporate store within the last year.) The results were similar for the 1-mile and 3-mile choices of radius for demographic variables but the variable-mile radii tended to have slightly better explanatory power and were generally regarded by management to be more relevant. The resulting models explained between 60% and 70% of the variation in sales-related variables, and about 60% of the variation in sales growth. To indicate the directional effect and the marginal utility of a variable (i.e. the relative increase in residual variance that would occur if the variable were removed from the predictive model for a performance indicator), we present, in Table 2, the *t*-statistics for each coefficient. Further, we denote, with one asterisk, a coefficient for which a two-tailed test of significance is significant at the 0.1 level;

with two asterisks a coefficient significant at the 0.05 level; with three asterisks a coefficient significant at the 0.01 level; and with four asterisks a coefficient significant at the 0.005 level. The *R*-squared values for the parsimonious models are given in the last row of the first section of Table 2. The explanatory power of the parsimonious models was close to that of the models constructed with the statistically significant factor scores. The latter, however, required all the individual variables to be computed.

From the broad array of items for which data were gathered, a few key variables provided the bulk of explanatory power in the regression models for the chain's performance in each market area. From the first estimating equation for average monthly sales, for example, we observed that, *ceteris paribus*, corporate store sales were higher where there was a higher concentration of black population, a lower concentration of people of Hispanic origin, where the 3-to-1-mile gradient in black population was relatively high, where median income was relatively low, where the market was dense, where selling area was greater, where neighbourhood crime rates were perceived by employees to be relatively high, in stores that were established more than six months ago, where the store was free-standing (rather than in a strip mall), in markets where the average distance to competitors was greater, where savings relative to competitors on the basket of staples were greater, and where crime in the vicinity of the corporate store in comparison with crime in the vicinity of competitors' stores was relatively low. Each of these findings is consistent with the chain's market positioning as a limited assortment discount store with a preponderance of private brands and some national brands. Its primary target market is the value-conscious consumer who considers both quality and price in selecting retail outlet and merchandise. Note that the demographic characteristics, corporate store characteristics, competitors' store characteristics and relative competitive position all emerged as statistically significant (on the margin) in explaining gross sales.

A similar set of inferences result when reading the *t*-statistics for the coefficients in the equation for sales per square foot. The external appearance of the corporate store, relative to that of competitors' stores, appeared as an additional factor in determining sales per square foot. Sales per customer visit, however, responded differently to several factors. They tended (*ceteris paribus*) to be lower in areas with a concentrated minority population but higher if there was a steeper gradient in the percentage of the population who were minorities and lacked high-school education. Corporate stores with greater selling area, *ceteris paribus*, tended to have higher gross sales and higher sales per customer, but lower sales per square foot. Market penetration as measured, *ceteris paribus*, was less for corporate stores with higher concentrations of Hispanic population and where total estimated expenditures on food were estimated to be greater for the market area. Because the latter appears in the denominator of market penetration, its heavy influence may reflect errors in estimating total grocery sales for the market area, or unrecognized competition on the fringes of the market area. Market penetration was relatively greater where the corporate store was freestanding, where the number of competitors was fewer, where average distance to competitors was greater, where the size of a store relative to the average size of competitors was greater, and when visibility from the major transportation artery and ease of access to the corporate store, in comparison with those of competitors' stores, were relatively good.

Growth in sales, *ceteris paribus*, was lower in areas with higher concentrations of Hispanic population, in areas with higher percentage of families in poverty and in

Table 2 *t*-statistics for variables in the parsimonious models for store performance[1]

Variable	*Sales*	*Sales per square foot*	*Sales per customer*	*Market penetration*	*Sales growth*	*Profit*
Demographic Variables						
Percent black	+2.23**		–2.42**			
Percent Hispanic	–2.14**	–2.99***	–3.56****	–4.44****	–3.25****	–2.20**
Gradient in percent black	+4.21****		+5.34****		–2.84***	
Gradient in percent Hispanic					+3.47****	
Estimated expenditures on food				–10.53****		
Percentage of families in poverty				+5.90****	–3.42****	–2.29**
Median household income	–2.32**	–1.74*				–4.95****
Percent without high school		+1.77*				
Median residential property value					–5.56****	
Gradient in food expenditures		–1.80*				
Gradient in percent poor		+2.39**				
Gradient in percent without high school			+5.07****			
Gradient in 1997 population		+2.25**				+2.50**
Dense market indicator	3.52****	+3.22****				+2.16**
Corporate store characteristics						
Weekly hours of operation		+1.67*				
Selling area	+3.11****	–4.91****	+2.62***			
Employee perception of crime rate	+3.24****	+4.62****	+2.96****			+2.75***
Fresh meat sold					+3.80****	–3.64****
Fresh meat added in 98					+2.98****	
Store opened within last 6 months	–2.32**	–2.76***	–2.28**		+7.26****	–3.97****
Store freestanding	+1.64*			+2.66***		+3.26***
Competitor characteristics						
Inverse of total competitors				+4.95****		
Average distance to competitors	+3.20****	+3.53****	+2.00**	+2.59**	+2.16**	+2.61**
Relative competitive position						
Average savings	+3.33****	+3.21****				+3.87****
Relative size				+2.48**		+2.97****
Relative external appearance		+2.66***	+2.51**			
Relative efficiency					+1.78*	
Relative security provided						+2.22**
Relative crime rate	–2.67***	–2.71***	–2.60**			
Relative visibility				+2.93****		
Relative ease of access				+2.77***		
R–squared	.64	.65	.61	.71	.51	.60

[1] Asterisks indicate two-tailed levels of statistical significance (* for 0.1; ** or 0.05; *** for 0.01; and ****for 0.005).

areas with lower housing values. Sales growth was higher where fresh meat was sold, where the store was established in the last 6 months, where fresh meat was added in the last fiscal year, where the average distance to the competitors is higher, and where efficiency of service, in comparison with competitors, was relatively good.

Finally, we examine the variables that were related to profit contribution. From the last column in Table 2, we observe that, *ceteris paribus*, profit contribution was higher where there was a low concentration of Hispanic people, where fewer families were in poverty, where median income was lower, where the gradient in population concentration was greater, and in dense markets. It was higher for free-standing stores, but lower in stores opened in the last six months and where fresh meat was sold (a recent innovation). On the competitive dimension, profits were higher, *ceteris paribus*, where the average distance to competitors was greater, where average savings relative to competitors' prices were greater, where store size relative to competitors was greater, where the appearance of stores relative to competitive stores was greater (actually, where store appearance was less overshadowed by a more upscale competitor), and where provisions for security were relatively better (again, where the gap relative to more upscale alternatives was less).

Tests against holdout sample

Data for stores in a dozen corporate markets were set aside for an "out-of-sample" test of the models' predictive utility. Selection of markets for the holdout sample was performed simply by selecting every 10th corporate store from the database which itself was organized by store identifier, and in turn, reflected geographical location. We thus had stores in the holdout sample that were widely scattered throughout the US. We applied the data for the stores in the respective markets to estimate the expected performance for stores in the holdout sample and regressed the actual performance measures against predicted performance measures. The resulting percentage of variation explained by the models and the corresponding levels of statistical significance are reported in Table 3.

For the models based on individual statistically significant variables in each category, the predictive power in the twelve holdout markets came predominantly from the demographic characteristics of the neighborhoods in which the stores were located. Neither store characteristics nor competitive characteristics alone were of statistically significant help in predicting performance. This could, of course, be due to the small holdout sample and to limited variability within the holdout sample with regard to those characteristics. Notable from the holdout results, however, was the significantly inferior performance of the models based on factor scores. We submit that the use of factor analysis to create orthogonal factors in place of individual variables (as a way of coping with multicollinearity) has two shortcomings. First is the possible loss of information that may be helpful for predictive purposes. Second is the need to continue to collect data on all the individual variables to compute the factor scores. Our findings reinforced these assertions.

Discussion of findings

Aspects of location are indeed important predictors of aggregate store performance. They are revealed through demographic statistics (including ethnicity) for the

Table 3 Proportion of variance in performance measures explained by different models for stores in the holdout sample[1]

Variables included	*Sales*	*Sales per square foot*	*Sales per customer*	*Market penetration*	*Sales growth*	*Profit*
All demographics	0.49***	0.32**	0.29**	0.69***	0.09	0.21*
3-mile demographics and gradients	0.41***	0.45***	0.27**	0.23*	0.21*	0.09
V-mile demographics and gradients	0.49***	0.41***	0.28***	0.77***	0.19*	0.10
Nonracial 3-mile demographics and gradients	0.35**	0.41**	0.29**	0.51***	0.48***	0.11
Nonracial *v*-mile demographics and gradients	0.40***	0.33**	0.32**	0.69***	0.40**	0.09
Store characteristics, v-mile demographics, and gradients	0.44***	0.55***	0.15	0.74***	0.05	0.22*
Store characteristics	0.10	0.10	0.01	0.01	0.03	0.17*
Competitive characteristics	0.07	0.00	0.04	0.25*	0.15	0.13
Store characteristics, v–mile demographic variables, gradient variables and competitive characteristics	0.46***	0.31**	0.27**	0.28**	0.05	0.31**
Parsimonious model using all variable categories	0.45***	0.44***	0.25**	0.65***	0.15	0.30**
Parsimonious models using non survey variables	0.38***	0.43***	0.32**	0.56***	0.00	0.20**
Demographic factor scores	0.26**	0.19*	0.36**	0.62***	0.28**	0.10
Demographic and store factor stores	0.20*	0.11	0.40**	0.67***	0.01	0.08
Demographic, store and competitive factor scores	0.09	0.35**	0.36**	0.28**	0.10	0.14

[1] Asterisks indicate two-tailed levels of statistical significance (* for 0.1; ** or 0.05; *** for 0.01; and ****for 0.005).

market area and also through the number and characteristics of competitive retail outlets. Gradients in demographic characteristics that express how the immediate market area compares with surrounding areas also contains significant information for assessing business potential. On spatial dimensions, sales, sales per customer, sales per square foot, market penetration and profit all tended to be greater (as expected) when the average distance to competitors' stores was greater. In addition to the spatial-locational aspects of competition, the importance of price positioning was revealed in the statistical significance of cost savings on a selection of staple products. Aesthetic appearance of stores in comparison with competitive stores, and comparative service levels surfaced in models constructed with a 0.2 level of statistical significance as an entry criterion, but they offered insufficient information on the margin to be retained in models that demanded a 0.1 level of statistical significance.

Sales levels, market penetration, sales growth and levels of profit are best explained by different sets of variables and respond in different directions to some of the same variables. The chain, for example, achieved greater market penetration in areas with concentrations of poor people, especially where the average distance to competitive facilities is large. Sales growth and profits, however, were lower in poorer areas. By fitting models to predict multiple performance measures, we can highlight the need for corporate decision-makers to consider the ways in which some objectives may be achieved at the expense of others.

Some additional findings emerged from the data collected at retail sites and the incidental conversations that occurred with store managers as surveys were executed. Most chains adopt uniform pricing strategies in a sales region. Although managers in several chains claimed to have some latitude in meeting local competition, the posted prices for individual products, even for those chains, were very common in a sales district (i.e. in areas served by a common wholesale outlet). Prices and promotions offered in poorer inner-city neighbourhoods were essentially the same as those offered by the same chain in affluent suburbs. Relative pricing by competitors can therefore be determined through efficient sampling rather than by surveying all stores. Chains do not employ risk-based pricing in high-crime urban areas, but they do vary the range of merchandise and special services.

Because of the target market for the limited-assortment discount chain that provided the data for this study, it was expected that accessibility to the store by public transit and availability of cross-shopping opportunities would be important elements in determining store performance. Neither of these variables, however, proved statistically significant as predictors of aggregate store performance. Store managers felt that only a small percentage of their shoppers used public transportation, except in a few highly urbanized areas. While the dense-market indicator may have served as a surrogate for public transit service in such cases, the private automobile is still by far the dominant form of transportation for food purchases in the United States – even among market segments least able to afford cars.

The spatial-locational aspects of competitive outlets (i.e. positioning relative to competitive outlets) emerge as a critically important variable with predictable effects on each of the six measures of store performance. Thus the lore of the importance of store location is supported by the findings presented here. Equally important are demographic variables, but the directional effects of some demographic variables (such as racial characteristics) depend on the characteristics of the product or service under consideration. They would also depend on the

market niche of the retail chain. Other characteristics shown in previous studies (Tantiwong *et al.* 1985; Mitchell and Kiral 1999) to affect the quality of the shopping experience for individual products were shown to have measurable effects on aggregate store performance as well.

Interestingly, the quality of merchandise and its presentation and the internal and external store ambience (relative to the average for competitors) emerged, in this study, as more important than the variety of goods and services in determining store performance. This may be specific to retailers that deliberately offer a limited selection of merchandise at discount prices. It may also be due to differences in the characteristics of the specific competitors that the firm faced in the different regions of the country. Nevertheless, stores with limited selling area face tradeoffs among quantity and variety of goods, appealing presentation of merchandise, and convenient circulation for customers. Careful experimentation to examine the effects of these tradeoffs for different types of merchandise seems warranted.

Managerial implications and conclusion

Returning to the issues of site evaluation for new stores, assessment of performance relative to potential at existing stores, and the collection of information for reviewing marketing strategies, we can offer several thoughts from our findings. Consider first the evaluation of new retail sites for a growing enterprise. We suggest that the analysis be conducted in a hierarchical fashion, with the intensity of the data collection effort increasing at each stage.

The evaluation of a potential site for a new facility would well begin with the collection of a few key demographic variables. Estimates of sales potential would be produced by applying the data to models that have been found to explain sales, sales per unit area and market penetration in existing markets. Population size, household income and ethnicity were the key elements for explaining store performance in the case of our discount grocery chain. Our findings reinforce the assertions of Fotheringham (1988) and Richards (1984) in that regard, and we found similar variables to be critical. Estimates of expenditures for grocery products, which were purchased at considerable cost from a commercial vendor of marketing data, contributed little additional information in our case, perhaps because those estimates were themselves based on the same demographic data and perhaps also because shoppers at the discount grocery chain tend to spend less per visit than shoppers at the major grocery chains.

Next in the hierarchy of information would be the spatial distribution of competitors and their general characteristics (market niche). The key variables at this level of analysis for the discount grocery store were the number of competitors who serve customers in the store's catchment area, the stores' service capacities, and the travel distances involved – exactly the items that are used in the various refinements of Huff's (1963) gravity model. Formal models of various types can be used to help determine the business potential for a new facility in a location with a particular combination of demographic characteristics and competitive outlets. For forecasting purposes, one could use various combinations of regression models, multinomial logistic models, MCI models, constrained allocation models, and perhaps neural networks. Gruca and Klemz (1998), for example, have found that neural networks performed relatively well in next-period forecasting of market shares of coffee brands, considering product pricing and

promotional mixes. Agrawal and Schorling (1996) suggest that neural networks would be useful in forecasting brand shares of grocery products if large amounts of scanner data were available. Like many researchers, Beaumont (1987) sees advantages in combining the results from several complementary approaches when performing studies for locating stores. From our own experience, we see the spatial-interaction (gravity or MCI) models as most useful for assessing the impact of the addition or deletion of service facilities in a specific market area. Regression models and time series models seem more effective for explaining or predicting sales performance in established markets and in making cross-sectional comparisons of performance among markets. With the power of a laptop computer alone and economical geographic information systems, one can quite easily test for consistency of assessment between regression-based approaches and gravity-based approaches based on the same data. Also, the outputs from gravity-based approaches may be used as inputs to a complementary regression model. Neural networks may provide another useful framework for combining the demographic data, spatial information and store-specific information.

Sites that pass muster at the first two stages of analysis (and existing stores) may be scrutinized further with detailed store surveys. The store surveys provide further information about the specific characteristics of stores and their environments. They also afford general assessments of product assortment, presentation, pricing and service-factors that have been demonstrated by Guerts and Whitlark (1993), Mulhern and Williams (1994) and Mitchell and Kiral (1999) to affect the choice of individual products and sales outlets. Using a semantic differential, a person can effectively describe relative visibility, ease of access, external appearance, general service efficiency, variety of merchandise and quality of merchandise. With a selection of benchmark products, price comparisons can also be made. While price comparisons for a market basket do take considerable time, we found in our study that stores in the same chain in the same region set prices consistently. We could therefore have chosen a single store from the chain to represent its prices in the region. Also, stores that were a similar size in the same chain had similar service characteristics. The name of the store itself carries a characteristic image with which representative values of the various variables seemed to be associated. The common characteristics of stores in a chain can therefore be exploited to create estimates of store-specific variables without a full-scale surveying effort, especially in preliminary analysis of a marketing area. These systematic patterns may contribute to the ability of senior managers (mentioned by Clarke *et al.* 2000) to develop an intuitive sense of the business potential at a site on the basis of what appears to be superficial information about the population and competition.

Although we have emphasized the selection of sites for new stores in our presentation, the models are also useful for reviewing the performance of existing stores (e.g. the stores comprising the analytical database). For this purpose, one may focus on the magnitudes of the residuals (differences between actual measure of performance and expected measure of performance as determined from the regression equation). Stores having positive residuals can be grouped and compared with stores having negative residuals in pursuit of common characteristics that may point to paths for improvement. Indeed, as managers are presented with a list of outlets ranked according to the differences between their actual and expected performance levels, they often identify factors that may account for the results – some of which can be remedied through managerial action. Stores with persistently

negative residuals need careful evaluation and, in the extreme, may be candidates for closure. Considering that our models explained, at best, 80% of the variation in sales in the markets used to calibrate the models and 50% when used for prediction in the holdout sample, there are other factors that are very influential in determining ultimate store performance. Grouping of stores by residual deviations for the various performance measures can help to uncover them.

Note that the variables which emerged in the first two levels of the analytical hierarchy for discount grocery stores were few in number, objective in nature and quite readily available. In the US and the UK, there are many sources of demographic, economic, psychographic and life-style data which can potentially help in the selection of a retail site, development of marketing strategy, and assessment of store performance. Because the value of information may be dependent upon the product or service and the marketing niche of the store, similar investigations in other types of retail enterprises would be warranted.

Managerial knowledge and intuition is an important element of the site-evaluation process, even when considerable reliance is placed on quantitative models. Subjective judgements of managers are involved in the identification of potential retail sites for which the demographic and spatial-locational information is gathered in the first place. The knowledge of experienced managers and marketing research staff is incorporated in the design of survey instruments to capture information that is particularly relevant to the product or service under consideration. They also play a critical role as the resulting statistical assessments are interpreted and choices are narrowed. Our findings confirm, to some degree, the suggestion of Blattberg and Hoch (1990) that forecasting would best consider data-based estimates tempered by managerial judgement (or vice versa). The subjective assessments of factors deemed important by managers did contribute marginal explanatory power when estimating store performance on several dimensions. We would encourage further research to determine the best ways of integrating managerial knowledge with data-based forecasts.

Acknowledegments

Daniel Eng, William Jackson and Delle Jackson did the vast majority of field surveys. Ajay Shah and Reza Mubeen drew lists of competitors, prepared maps of the market areas to be surveyed, and created the computerized database of survey information. Our colleague Haim Mano offered helpful comments on the first draft.

References

Achabal, D.D., Gorr, W.L. and **Mahajan, V.** (1982) 'Multiloc: a multiple store location decision model', *Journal of Retailing*, 58(2): 5–25.

Agrawal, D. and **Schorling, C.** (1966) 'Market share forecasting: an empirical comparison of artificial neural networks and multinomial logit model', *Journal of Retailing*, 72(4): 383–407.

Beaumont, J.R. (1987) 'Retail location analysis: some management perspectives', *International Journal of Retailing*, 2(3): 22–35.

Blattberg, R.C. and **Hoch, S.J.** (1990) Database models and managerial intuition: 50% model + 50% manager', *Management Science*, 36(8): 887–99.

Bodkin, C.D. and **Lord, D.J.** (1997) 'Attraction of power shopping centres', *The*

International Review of Retail, Distribution and Consumer Research, 7(2): 93–108.

Bucklin, L. (1971) 'Retail gravity models and consumer choice: a theoretical and empirical critique', *Economic Geography*, 47(4): 489–96.

Cadeau, J.M. (1999) 'Category size and assortment in US macro supermarkets', *The International Review of Retail, Distribution and Consumer Research*, 9(4): 367–77.

Clarke, I., Horita, M. and **Mackaness, W.** (2000) 'The spatial knowledge of retail decision makers: capturing and interpreting group insight using a composite cognitive map', *The International Review of Retail, Distribution and Consumer Research*, 10(3): 265–85.

Cooper, L.G. and **Nakanishi, M.** (1988) *Market Share Analysis: Evaluating Competitive Marketing Effectiveness*, Boston, MA: Kluwer Academic Publishers.

Epstein, B.J. (1984) 'Market appraisals', in Davies, R.L. and Rogers, D.S, (eds) *Store Location and Store Assessment Research*, Chichester: Wiley.

Fotheringham, A.S. (1988) 'Market share analysis techniques: a review and illustration of current US practice', in Wrigley, N. *Store Choice, Store Location and Market Analysis*, New York: Routledge.

Ghosh, A. (1984) 'Parameter nonstationarity in retail choice models', *Journal of Business Research*, 12(4): 425–36.

Ghosh, A. and **McLafferty, S.L.** (1987) *Location Strategies for Retail and Service Firms*, Lexington, MA: D.C. Heath and Co.

Goodwin, D.R. and **McElwee, R.E.** (1999) 'Grocery shopping and an ageing population: research note', *The International Review of Retail, Distribution and Consumer Research*, 9(4): 403–9.

Gruca, S. and **Klemz, B.R.** (1998) 'Using neural networks to identify competitive market structures from aggregate response data', *OMEGA,* 26(1): 49–62.

Guerts, M.D. and **Whitlark, D.** (1993) 'Forecasting market share', *Journal of Business Forecasting*, 11(4): 17–22.

Huff, D.L. (1963) 'A probabilistic analysis of shopping center trade areas', *Land Economics*, 39: 137–76.

Koelemeiger, K. and **Oppewal, H.** (1999) 'Assessing the effects of assortment and ambience: a choice experimental approach', *Journal of Retailing*, 75(3): 319–45.

Korgaonkar, P.K., Lund, D. and **Price, B.** (1985) 'A structural-equations approach toward examination of store attitude and store patronage behavior', *Journal of Retailing*, 61(2): 39–60.

Lundsten, L.L. (1978) 'Market share forecasting for banking offices', Ann Arbor, MI: UMI Research Press.

Mahajan, V., Jain, A.K. and **Ratchford, B.T.** (1982) 'Use of binary variables in the multiplicative competitive interactive choice model', *Journal of Consumer Research* (5): 210–15.

Mahajan, V., Sharma, S. and **Srinivas, D.** (1985) 'An application of portfolio analysis for identifying attractive retail locations', *Journal of Retailing*, 61(4): 19–34.

Mitchell, V.-W. and **Kiral, H.R.** (1999) 'Risk positioning of UK grocery multiple retailers', *The International Review of Retail, Distribution and Consumer Research*, 9(1): 17–39.

Moutinho, L., Curry, B. and **Davies, F.** (1993) 'Comparative computer approaches to multi-outlet retail site decisions', *Services Industries Journal*, 13(4): 201–20.

Mulhern, F.J. and **Williams, J.D.** (1994) 'A comparative analysis of shopping behavior in Hispanic and non-Hispanic market areas', *Journal of Retailing*, 70(3): 231–51.

Nakanishi, M. and **Cooper, L.G.** (1974) 'Parameter estimates for multiplicative competitive interaction models – least squares approach', *Journal of Marketing Research*, 11(August): 303–11.

Nakanishi, M. and **Cooper, L.G.** (1982) 'Simplified estimation procedures for MCI models', *Marketing Science*, 1: 314–22.

Penny, N.J. and **Broom, D.** (1988) 'The Tesco approach to store location', in Wrigley, N. (ed.) *Store Choice, Store Location and Market Analysis*, New York: Routledge.

Richards, D. (1984) 'General sources of information', in Davies, R.L. and Rogers, D.S. (eds) *Store Location and Store Assessment Research*, Chichester: Wiley.

Simonson, I. (1999) 'The effect of product assortment on buyer preferences', *Journal of Retailing*, 75(3): 347–70.

Smith, L.D. and **Moses, S.L.** (1996) 'Strategic planning of transportation services for petroleum products: an application of capacitated gravity models', *European Journal of Operational Research*, 88: 215–30.

Tantiwong, D. and **Wilton, P.C.** (1985) 'Understanding food store preferences among the elderly using hybrid conjoint measurement models', *Journal of Retailing*, 61(4): 35–64.

Tyagi, R. and **Das, C.** (1997) 'A methodology for cost versus service trade-offs in wholesale location-distribution using mathematical programming and analytical hierarchical process', *Journal of Business Logistics*, 18(2): 77–99.

Zeller, R.E., Achabal, D.D. and **Brown, L.A.** (1980) 'Market penetration and locational conflict in franchise systems', *Decision Sciences*, 11(1): 58–80.