Advances in methods to support store location and design decisions

Hunneman, Auke

Publication date:
2011

Chapter 5

Conclusions and Further Research

5.1 Introduction

Retailing literature provides compelling evidence that spatial convenience, that is, the proximity of stores to consumers, is a primary driver of store choice, closely followed by assortment and prices (Briesch, Chintagunta, and Fox 2009; Fox and Sethuraman 2006; González-Benito, Bustos-Reyes, and Muñoz-Gallego 2007). Although this finding underlines the strategic importance of location decisions, it does not imply at all that retailers can just ignore other marketing variables when establishing a store at a particular location. Consumers are more likely to visit a store (format) that matches their consumer profile for example, so store managers can differentiate along dimensions other than space (Inman, Shankar, and Ferraro 2004). The mere presence of competitors in a particular market may force (nearby) stores to adopt different formats, through which they can mitigate the effects of spatial competition. Finally, retail properties themselves come at a cost that also determine local pricing strategies and margins (Fox and Sethuraman 2006). Thus, location is an important marketing variable that should be considered not in isolation but rather as an important facet of the retailer's overall strategy.

The main objective of this dissertation has been to develop new methodologies for store location evaluation and choice. All the proposed methodologies explain store sales at the zip code level and account for substantial heterogeneity in consumer characteristics. The models include the spatial distribution of consumer demographics; previous studies mainly have used aggregate measures for these variables. In this concluding chapter, we summarize the key findings of the research that constitutes this dissertation. We also answer the research questions formulated in Chapter 1 and identify new ones in need of additional scientific inquiry. These avenues for research are discussed further in the final section of this chapter.

5.2 Main Findings

Many retail industries today are dominated by chain outlets; the top 50 supermarket chains in the United States for example operate an average of 378 stores each (Gauri, Pauler, and Trivedi 2009). Location decisions for such large store networks are complicated by the vast number of interactions between individual stores and their powerful performance implications for the whole network. Chapter 2 therefore proposes an optimization model to support location decisions for chain stores, which can determine the optimal number of stores, store locations, and store sizes for a given region simultaneously. Because the model uses the mixed integer linear programming (MILP) paradigm, instances of considerable sizes (i.e., potential store locations) can be solved effectively, which fosters its implementation. The proposed model also presumes that consumers choose among alternative stores on the basis of travel distance to the store and store size—a more realistic assumption than that of the so-called full-capture model that only considers distance. In an empirical application to health clubs in the Rotterdam area, we show that the model replicates the findings of the full-capture model if the store size sensitivity parameter is set to 0. We further demonstrate that the proposed model generates better solutions than a model that first determines optimal store locations and then the store sizes for each location. The difference in profit levels between these two approaches depends (non-monotonically) on the size of the store size sensitivity parameter.

In Chapter 3, we consider the situation in which a retail chain wants to add a new store to its existing network of stores. We have developed a model that supports evaluations of the sales impacts of changes in a store's retail environment, location changes, and new store openings. Rather than modeling total store sales, we split this variable into its constituent parts and develop separate models for loyalty card penetration, average number of visits, and average expenditures per visit. Previous research reveals that the effect of predictor variables may differ across these sales components (Pan and Zinkhan 2006; Van Heerde and Bijmolt 2005), so a model evaluating the effect of each predictor variable on each sales component separately is likely to offer richer insights than a single model for just sales. We find accordingly that the effect of our explanatory variables differs across the loyalty card penetration rate, the average number of visits, and the average amount spent. For example, distance to the store negatively affects the penetration rate and the average number of visits but has a positive relationship with average expenditures per visit. The results further provide evidence of spatial dependence across the observations for each sales component, due to the unobserved similarities of consumers living in close proximity. We evaluate the usefulness of the decomposition framework in an empirical application involving 28 stores of a Dutch clothing retailer. The findings show that the decomposition model achieves high predictive performance and is able to forecast the performance of new stores reasonably well. The model accounting for spatial autocorrelation also performs better than the models ignoring spatial dependence and can identify areas in which store performance can be improved, according to a comparison of the realized sales figures with those predicted by a model for a store's trade area. Another application evaluates the sales impacts of changes in predictor variables such as the relative sizes of individual store departments. The results of these analyses can help retailers establish an optimal assortment for each store location.

In Chapter 4, we extend these approaches by considering a store's overall performance as the sum of the performance of individual store departments. Department-level sales likely vary by space, due to differences in their attractiveness to local consumer groups. The proposed model allows for spatial heterogeneity in consumer preferences by modeling the sales shares of each individual department at the zip code level, as a function of the characteristics of that local environment. The findings indicate that the sales levels of individual store departments are affected differentially by location variables, which offers room for improvements in store performance if the retailers were to adjust each store's assortment to the local environment. Total sales levels are higher in zip codes near focal stores with a potentially large market and in which the number of competitors is high. We also find evidence that store managers allocate more space to relatively stronger departments. Specifically, departments with better past performance (i.e., last year's sales) tend to occupy a larger proportion of a store's floor space, at the expense of other departments. Therefore, (relative) department sizes should be considered endogenously; ignoring this reverse causality likely leads to overestimations of the effects of department size on sales. The model in Chapter 4 also can predict potential sales of new stores, which makes it a useful tool for store location evaluation. We apply this model to evaluate the performance implications of changes in the assortment composition of new stores, an application never done before (Campo et al. 2000).

The combined results of the three studies presented in this dissertation add to store location literature in several important ways. First, by using disaggregated data at the zip code level, our proposed models account for more heterogeneity in consumer characteristics than do currently available models, independent of whether this heterogeneity occurs in (researcher-) observed or unobserved variables. Second, to forecast (future) values for each sales component in a particular zip code, we borrow information from neighboring zip codes. Taking this spatial dependence between zip codes in close proximity into account leads to better predictions, which enhances the usefulness of the proposed models for store location and evaluation decisions. Third, rather than evaluating the impact of drivers of store performance on a store's total sales level, we disentangle the effects of each predictor variable for various components of store sales and thereby offer richer insights than a single model for just sales would. The predictor variables have different effects on each sales component, such as sales of individual store departments. In such a situation, a retailer can enhance the performance of its individual stores by tailoring the assortment of each store to the characteristics of the local markets in which they operate.

From a practical perspective, we demonstrate applications of the proposed models to (1) identify and evaluate new store locations; (2) measure the sales impact of changes in store location, store assortment design, and the retail environment; and (3) assess the (relative) performance of existing stores. As we noted in Chapter 1, we recommend a successive application of the proposed models, in the same order as they appear in this dissertation.

5.3 Limitations and Further Research

5.3.1 Endogeneity of Market Structure

A major limitation of the empirical studies in this dissertation is that we do not always fully correct for the endogeneity of observed market structures, which potentially leads us to assess the effects of some predictor variables on store performance incorrectly. Market structure is a broad concept that entails the number, size, and distribution of buyers and sellers in a market (American Marketing Association 2010). If we contrast the elements of this definition with market structure variables, we would conclude that our featured elements are rather limited. For example, the model in Chapter 3 only considers the number of competitors as an endogenous variable, and competition is treated strictly exogenously in the other chapters. Yet recent studies using observed store location patterns across markets in the U.S. retail industry show that market entry and location decisions depend heavily on the (actual and anticipated) positioning of competitors (Hansen and Singh 2009; Seim 2006; Thomadsen 2007; Zhu, Singh, and Manuszak 2009). Therefore, though retailers prefer to locate stores closer to greater sources of demand, the (anticipated) presence of (direct) competitors may prevent them from doing so. Moreover, rivals in close proximity may weather competition better by differentiating themselves, whether geographically or demographically. If these potential interdependencies among firms' decisions are ignored, sales forecasts might be biased. Consequently, misleading conclusions could result about the effect of competition on store revenues. In such a situation, a retail manager might wrongly decide or not to open a new store in a particular market because he or she underestimates or overestimates the potential detrimental effects of competition on potential performance.

Therefore, further research should include additional information about competitors, including their physical store locations, other store attributes, and marketing mixes. Collecting these data could pose significant challenges, because information about competitors' marketing activities is likely confidential or simply not available, and the number of competitors tends to be large. However, if we succeed in obtaining more information about competitors, we might build on a small but growing literature stream that allows for strategic interactions in firms' decision making.

Seim (2006) and Zhu and Singh (2009) consider a situation in which firms not only decide whether to enter a particular market but also, if they enter, where to locate the new stores. The econometric specification of these models relies on a game-theoretic model that assumes a firm's (latent) profit in a particular location depends on distance from that location to the firm's nearest store and the relative proximity of competitive stores. The model developed by Zhu and Singh (2009) also can be extended to include the (most appropriate) store format for each location, similar to the tailoring of assortment shares to local conditions in Chapter 4. They use a random coefficients model that allows for asymmetric competitive effects due to firm-specific differences (Ailawadi et al. 2009). Therefore, these models meet some of the limitations of the models developed herein, though it should be noted that they suffer their own shortcomings. Possible extensions could include location-specific unobservable variables (such as traffic patterns and the presence of employers, see, e.g, Duan and Mela 2009); dynamics in strategic decision making, such as market entry and exit; or interactions across markets (Bronnenberg and Mahajan 2001).

5.3.2 Market Area Delineations

Another important avenue for research is the delineation of market areas. In this dissertation, we have defined market areas as zip codes that encompass 85 percent of a store's sales. However, because sales depend on the retailer's decision to open a store at a particular location, this trade area definition cannot be treated as an exogenous explanatory variable. We need another way to determine trade area boundaries, one not based on the amount of sales that a location is able to generate. Although the problem of determining a store's trade area has been addressed from many perspectives, it still seems promising to model an individual consumer's store choice decisions with a spatial choice model that allows for heterogeneous consumer preferences. This model can then predict the probability that consumers from a certain location visit a particular store while incorporating (spatial) differences in consumer demographics, store attributes, competition, and travel distances to stores. If modelers were to supplement such information with a market's geographic population density surface, they could infer local demand. Donthu (1991) and Donthu and Rust (1989) use kernel density estimation methods to find a geographic population distribution from a random sample of addresses, but researchers could also use an interpolation method such as spatial kriging (Duan and Mela 2009). If they combine store patronage predictions with the inferred population densities, they might calculate the market shares for each individual store. Furthermore, because a competitor opening a new store in a particular market changes the explanatory variables of the store choice model, it is possible to evaluate the impact of such competition on the market shares of an individual store. One of the main advantages of this approach is that it does not just dismiss location information by aggregation into subareas, such as zip codes.

It is interesting to note that kernel density estimation also has been applied to determine consumer ideal points in perceptual maps (Rust and Donthu 1988). These points represent ideal combinations of product or store attributes from a consumer's perspective and are therefore attractive sites for firms introducing a new product. For retail stores, we might define these sites along several dimensions, such as the travel distances between households and stores and store profiles, which can be defined as factors of consumer demographics and/or utility distances (Van Dijk et al. 2004). If researchers were to contrast the two (perceptual) maps obtained from such an analysis against the actual store locations, they could evaluate the extent of competition along each dimension. An outcome therefore might be that adjacent stores do not compete very much, because they have different store profiles, whereas distant stores compete more intensely because they target the same clientele. This possibility is even greater if stores in close proximity target the same customer groups with different portfolios. In this scenario, the stores might be complementary, such that each individual store benefits from small interstore distances (Arentze, Oppewal, and Timmermans 2005; Dellaert et al. 1998; Popkowski Leszczyc, Sinha, and Sahgal 2004).

In summary, the models in this dissertation might benefit from efforts to make more variables endogenous. However, we concur with Shugan (2004): Models are primarily developed to assist decision makers. In some situations, it might be better not to model particular variables, such as the availability of locations due to zoning regulations. Adding these constraints can make models unnecessarily complicated and therefore detract from their usefulness for real-world marketing practice.