Advances in methods to support store location and design decisions
Hunneman, Auke

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2011

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 4

Evaluating Store Location and Assortment Design Based on Spatial Heterogeneity in Sales Potential

4.1 Introduction

Store location and assortment composition are important strategic decisions for many retailing firms. For example, Walmart announced that it would add about 3.4 million m² globally in 2011, which will demand capital expenditures of $13–15 billion (Progressive Grocer 2009). The growth of retailers depends largely on their selection of geographical markets and opening of new stores. Retail chains that find the best match between the positioning of their stores and characteristics of the local market are the most likely to succeed (González-Benito, Bustos-Reyes, and Muñoz-Gallego 2007). From a consumer perspective, convenient locations and product assortment drive store choice (Briesch, Chintagunta, and Fox 2009). Therefore, the evaluation of (potential) sites should consider store location and assortment composition together.

The best regions for opening new stores are those that generate the highest demand or sales (Levy and Weitz 2004). The problem, however, is that (potential) sales are not readily observed and do not necessarily match with observed population density (Duan and Mela 2009; Garber et al. 2004). Therefore, retailers use population characteristics associated with local market potential and buying power to infer potential sales from candidate sites (e.g., Kumar and Karande 2000; Putler, Kalyanam, and Hodges 1996). Beyond their impact on store choice (Pan and Zinkhan 2006), location characteristics, such as the geodemographic profile of customers and

---

1 This chapter is based on a working paper with the same title co-authored by Tammo H.A. Bijmolt and J. Paul Elhorst.
the presence of competitive stores, affect the relative attractiveness of different product categories and therefore consumer spending patterns (Inman, Shankar, and Ferraro 2004). Retailers that customize their product assortment at the store level according to location characteristics can dramatically increase their profits (e.g., Campo et al. 2000). A logical extension of retail location models therefore is to work to improve store performance by tailoring the assortment composition to local conditions. We therefore develop a methodological model that supports store location and assortment composition for new retail store locations.

This chapter contributes to extant literature in several ways. Our store location model can be used to determine the performance implications of changes in the store’s assortment composition for each proposed site. Although other studies have considered location choice in combination with other marketing mix variables before, we combine outlet location and assortment composition for the first time. We also relate department-level store sales to store, competitor, and consumer characteristics and thus provide rich insights into the drivers of department sales. By investigating consumer data at the zip code level, we allow for heterogeneity in consumer characteristics and preferences over space. We further account for unobserved spatial effects in department sales by including spatially autocorrelated error terms.

The remainder of this chapter is organized as follows: We start with a summary of the relevant literature in section 4.2, then introduce our models to explain total chain sales (4.3.1), department sales shares (4.3.2), and the relative sizes of each department (4.3.4). Section 4.4 elaborates on the estimation procedure for the attraction models, followed by a discussion of how these models can help predict each department’s sales share and its relative size among the store’s total floor space (section 4.5). We apply our modeling framework in an empirical setting (section 4.6), the results of which we discuss in section 4.7. Sections 4.8 and 4.9 suggest two potential applications of the proposed models, namely, the evaluation of new store
locations and the impact of changes in the assortment of each store. Finally, we present some conclusions, managerial implications, and directions for research.

4.2 Related Literature

Our work lies at the intersection of several research streams in marketing; we discuss three (see Table 4.1 for an overview). First, we note work on micromarketing, particularly that which shows that outlet location moderates the optimal assortment. We also position our work against spatial econometric models that take spatial dependence among observations into account. Finally, we discuss empirical economics literature, which simultaneously considers outlet location and marketing mix decisions.

4.2.1 Micromarketing and Shelf Space Allocation

Of growing interest to practitioners and academics alike is the possibility of exploiting spatial differences in category appeal by tailoring assortments to local needs. This example of micromarketing is the type of strategy adopted by retailers to tailor their marketing mix elements at the store level instead of following the same policy for every store in the chain (Montgomery 1997). Several factors are responsible for the widespread application of micromarketing, including the desire of retail managers to find new ways to differentiate themselves and lower costs at the same time (Campo et al. 2000; Desmet and Renaudin 1998; Grewal et al. 1999). Furthermore, the adoption of customer loyalty cards and the availability of scanner data have offered retailers greater possibilities for analyzing heterogeneity in consumer preferences and customizing their assortment accordingly.
<table>
<thead>
<tr>
<th>Article</th>
<th>Consumer Response Variable</th>
<th>Marketing Decision(s)</th>
<th>Explanatory Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel (a): Micromarketing</strong></td>
<td></td>
<td></td>
<td>Store</td>
</tr>
<tr>
<td>Grewal et al. (1999)</td>
<td>Efficiency</td>
<td>Assortment planning</td>
<td>x</td>
</tr>
<tr>
<td>Campo et al. (2000)</td>
<td>Shelf space elasticity</td>
<td>Shelf space allocation</td>
<td>x</td>
</tr>
<tr>
<td>Campo and Gijsbrechts (2004)</td>
<td>Shelf space elasticity</td>
<td>Shelf space allocation</td>
<td>x</td>
</tr>
<tr>
<td>Gijsbrechts, Campo, and Goossens (2003)</td>
<td>Flyer design elasticity</td>
<td>Feature promotions</td>
<td>x</td>
</tr>
<tr>
<td>Hoch et al. (1995)</td>
<td>Price elasticity</td>
<td>Pricing</td>
<td>x</td>
</tr>
<tr>
<td>Montgomery (1997)</td>
<td>Price elasticity</td>
<td>Pricing</td>
<td>x</td>
</tr>
<tr>
<td>Mulhern, Williams, and Leone (1998)</td>
<td>Price elasticity</td>
<td>Pricing</td>
<td>x</td>
</tr>
<tr>
<td><strong>Panel (b) Spatial models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bronnenberg and Mahajan (2001)</td>
<td>Price promotion elasticity</td>
<td>Price promotions</td>
<td>x</td>
</tr>
<tr>
<td>Van Dijk et al. (2004)</td>
<td>Shelf space elasticity</td>
<td>Shelf space allocation</td>
<td>x</td>
</tr>
<tr>
<td><strong>Panel (c): Outlet location &amp; Micromarketing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan, Padmanabhan, and Seetharaman (2007)</td>
<td>Consumer utility</td>
<td>Outlet location &amp; pricing</td>
<td>x</td>
</tr>
<tr>
<td>Duan and Mela (2009)</td>
<td>Consumer utility</td>
<td>Outlet location &amp; pricing</td>
<td>x</td>
</tr>
<tr>
<td>Thomadsen (2007)</td>
<td>Consumer utility</td>
<td>Outlet location &amp; pricing</td>
<td>x</td>
</tr>
<tr>
<td>This study</td>
<td>Shelf space elasticity</td>
<td>Outlet location &amp; assortment composition</td>
<td>x</td>
</tr>
</tbody>
</table>
Many micromarketing studies have concentrated on explaining differences in consumers’ reactions to marketing activities across stores. They typically explain heterogeneity in consumer responses, indicated by differences in price (e.g., Hoch et al. 1995; Kamakura and Kang 2007; Mulhern, Williams, and Leone 1998) or shelf space elasticities (e.g., Desmet and Renaudin 1998), with a second-stage analysis that relates differences in response variables across stores to store and trade area characteristics. Although these studies show that consumers may respond differently to marketing activities across stores, they do not explain why these differences exist or how retailers might exploit them.

Campo and Gijsbrechts (2004) and Campo et al. (2000) use normative models to find the optimal shelf space shares for each product category in a particular outlet, according to local differences in consumer preferences and competition. In particular, Campo et al. (2000) demonstrate that category attractiveness depends on store and trade area characteristics; if more shelf space is allocated to locally appealing categories, chain profits can be improved considerably. Moreover, Campo and Gijsbrechts (2004) show that micro-marketing strategies depend on the store format, so the optimal shelf space allocation should be differentiated across formats.

This stream of literature (e.g., Campo et al. 2000; Chen et al. 1999) also suggests three ways outlet location and micromarketing may drive store sales. First, location factors such as local buying power and the number of inhabitants of the store’s trade area can lead to a direct shift in store sales. These direct effects influence all product categories and are fairly well documented in retailing literature (e.g., Levy and Weitz 2004; McGoldrick 1990). Second, location factors can have differential impacts on the local attractiveness of different product categories. In particular, the space assigned to locally attractive categories may prompt a store draw effect that attracts new customers to the store, as well as prompt current customers to spend more. Third, local differences in category appeal determine
shopping basket compositions, such that product categories with strong local appeal represent a larger proportion of store sales, as do their complements, whereas substitutes for locally appealing products likely sell less. In addition, category sales may vary from one location to another as a result of differences in category-specific competition.

4.2.2 Spatial Models of Demand

Spatial econometric models apply to a broad range of marketing problems (for overviews, see Bradlow et al. 2005; Bronnenberg 2005). An important advantage of these models is their potential to account for unobserved firm behavior by combining data from multiple markets. For example, Bronnenberg and Mahajan (2001) show that retailers use the current demand levels of three national brands, unobserved to the researcher, to set their advertising and promotion expenditures for these brands in a particular market. If a retailer decides to invest more heavily in markets where the brand (or product category) is already a large share player, we expect a positive correlation between expenditures on marketing variables and (anticipated) sales. Bronnenberg and Mahajan (2001) also effectively capture unobserved retailer behavior by assuming a joint spatial dependence of marketing variables and sales. Using a spatial structure in the error terms based on geographic distances among stores, they estimate more realistic (i.e., smaller) absolute price promotion elasticities than can a model that assumes the predictor variables are truly endogenous. Disentangling simultaneous effects becomes more complicated for variables that change slowly over time, as is typically the case for the amount of shelf space. In these situations, we can better exploit differences in the amount of shelf space attributed to a particular product across stores. Van Dijk et al. (2004) use an approach similar to that of Bronnenberg and Mahajan (2001) to obtain shelf space elasticities for five Dutch brands in the shampoo category. However, instead of using a spatial structure based on geographic proximity, they use similarities
between store profiles to obtain reliable shelf space elasticities. They argue that if managers make decisions according to store profiles, it is more informative to use a profile-based similarity measure than geometric distance. Geometric distances are especially inadequate in situations in which stores in close proximity have more dissimilar client groups than stores located farther apart. The results of these studies thus indicate that spatial econometric models can efficiently capture geographical variations in demand and supply side factors, even if these factors are not observed.

4.2.3 Structural Models of Outlet Location

The simultaneous consideration of location and marketing mix decisions has only recently received research attention, such as by Chan, Padmanabhan, and Seetharaman (2007), Duan and Mela (2009), and Thomadsen (2007), who determine equilibrium prices or sales conditioned on outlet location and capacity. Most work in this area belongs to a developing research stream in marketing: empirical economics (Chintagunta et al. 2006). Although we adopt a reduced form approach, we note the relevant empirical economics literature here.

Specifically, Chan, Padmanabhan, and Seetharaman (2007) develop an econometric model of both geographic location and price competition among gasoline stations in Singapore. Their location model exploits the observed spatial distribution of geodemographic variables to infer potential gasoline demand at each local market. The location decision is an optimization problem, in which the Singapore government determines optimal locations from a social welfare perspective. Therefore, the total sum of traveling distances to different store locations across all consumers gets minimized, instead of some firm-specific measure of interest, such as profits or sales.

In a similar vein, Thomadsen (2007) parameterizes spatial demand as a function of observed geographic characteristics, prices, and the travel distances of consumers to stores. He uses demand estimates for each location to identify the
most profitable outlet locations for fast-food chains with asymmetric competitive strengths. The results indicate that pricing and location decisions are interrelated, because optimal outlet locations depend on (price) competitive intensity and the size of the market at a particular location. However, firms may adjust their prices in response to the geographic layout of the market. The results indicate that retailers that want to open a new store should evaluate candidate sites carefully according to their potential sales, as well as the intensity of competition.

These studies by Chan, Padmanabhan, and Seetharaman (2007) and Thomadsen (2007) use observed spatial differences in consumer characteristics and consumers’ distances to the store to infer spatial demand. Duan and Mela (2009) supplement such observed spatial demand factors with spatially correlated unobserved demand effects. They also augment spatial statistics with a structural model of pricing, which they use to simulate the effect of changes in outlet locations on equilibrium prices and profits.

In summary, this literature stream models consumer choice of outlet locations and shows that variations in demand across locations can be explained by product and outlet characteristics, competition, and the spatial and demographic distribution of consumers within a market. Yet outlet locations are considered endogenous, so they do thus not depend on (the retailer’s conjectures about) the location decisions of others. Related literature (Mazzeo 2001; Seim 2006; Zhu and Singh 2009) asserts that entry and location decisions depend on the potential profitability of a particular market and inferences about (future) competitor decisions. Zhu and Singh (2009) find, for example, that discount retailers prefer to locate stores as close as possible to (potentially) large markets, but that the threat of competition may prevent them from doing so. Differentiation, through the adoption of different store formats or assortment, may weaken the impact of such competition.
4.2.4 *This Study*

We address some important issues not covered by existing literature (Table 4.1). For example, though some traditional micromarketing papers note the impact of locational factors on assortment composition, they consider only the allocation of shelf space for existing stores. As suggested by Campo et al. (2000), the appropriate assortment for new stores also can be determined from their location profile, if known, which is the focus of our study. Furthermore, existing micromarketing literature addresses assortment composition at either the product category (e.g., Campo et al. 2000; Campo and Gijsbrechts 2004) or brand (e.g., Kamakura and Kang 2007; Montgomery 1997) level, ignoring store sales at the department level.

Similar to Duan and Mela (2009), we build on spatial modeling literature by using a model that accommodates unobserved spatial effects in store sales variables. We assume that zip codes in close proximity share unobserved characteristics, which may cause spatially correlated error terms. Failing to account for spatial error autocorrelation when it exists may cause inefficiency (Anselin 1988). Therefore, we adopt models with spatial autocorrelation to account for spatial dependencies in sales components across zip codes and stores. In addition to allowing for unobserved sources of store sales, our model uses a broad range of location characteristics, such as store, competitor, and consumer characteristics observed at the zip code level, to infer department-level sales.

Unlike prior research that considers both outlet location and marketing mix decisions (Chan, Padmanabhan, and Seetharaman 2007; Thomadsen 2007), we investigate assortment composition rather than pricing—the predominant topic thus far. Furthermore, most existing work is built on the premise of Bertrand competition among firms, such that equilibrium prices depend on differences in competitive intensity rather than variations in consumers’ price sensitivity across markets. These studies thereby assume that competitors offer nearly the same products, which is obviously not the case for many stores. It is therefore desirable to ascertain whether
differences in consumer preferences and competition across products lead to
different outcomes. We accordingly model store sales at the department level.

4.3 Model Specification

We define total chain sales as the sum of sales over all departments and zip codes in
the area in which the chain operates, equivalent to the product of a department’s
share of sales and the total amount of sales generated by the chain at that zip code:

\[ S_t = \sum_{m=1}^{M} \sum_{j=1}^{J} S_{mjt} = \sum_{m=1}^{M} \sum_{j=1}^{J} DSS_{mjt} \times S_{jt}, \]  

(4.1)

where \( j \) refers to zip codes (\( j = 1, \ldots, J \), such that \( J \) is the number of zip codes),
\( m \) denotes the department (\( m = 1, \ldots, M \), where \( M \) is the number of departments),
and \( t \) represents a given time period (\( t = 1, \ldots, T \), and \( T \) is the number of time
periods). Furthermore, \( S_t \) is chain-level sales to members at time \( t \), \( S_{mjt} \) refers to
sales of department \( m \) in zip code \( j \) at time \( t \), \( DSS_{mjt} \) is the sales share of
department \( m \) in zip code \( j \) at time \( t \), and \( S_{jt} \) represents the chain-level sales in
zip code \( j \) at time \( t \).

Customers signing up for loyalty programs must provide the retailer with their
addresses, so their subsequent purchases are registered by the system. We use this
information to obtain detailed insights about the mechanisms driving store sales.
Loyalty program members usually are responsible for a large proportion of total
sales (Singh, Hansen, and Blattberg 2006; Van Heerde and Bijmolt 2005); we
therefore limit our analysis to purchases by these loyalty card customers. By
gathering information about the residence of these members, we in turn determine
how department-level store sales are distributed geographically.

We develop models for both variables on the right-hand side of Equation 4.1,
\( DSS_{mjt} \) and \( S_{jt} \), to capture the different ways in which the outlet location may
affect store sales. Consider for example the effect of an increase in consumers’
buying power in a particular area, which is common to all departments and thus might lead to higher store sales but not necessarily changes in departments’ sales shares. This effect is captured by the model that explains the total amount of sales generated in a particular zip code. Other location characteristics, such as the number of children living in a particular area, also might differentially affect the attractiveness and sales shares of individual departments (e.g., children’s, women’s clothes), accounted for in the model by each department’s share of sales at a particular location. With this modeling approach, we also capture the different ways in which assortment composition affects store sales. A store’s total assortment drives its attractiveness to certain consumer groups and thus store choice, which generally then leads to higher overall sales (Briesch, Chintagunta, and Fox 2009; Chernev and Hamilton 2009). Moreover, changes in the assortment composition of a store may lead to relatively higher sales shares for categories/departments that constitute a larger proportion of a store’s assortment.

Leaving aside the two variables on the right-hand side of Equation 4.1, we recognize that if retailers allocate more floor space to departments that perform well in a particular store, the reverse effect may emerge. A department’s (past) sales may determine its (relative) size in the store, as a result of which department sizes should be considered endogenous. We therefore also develop a model to explain the (relative) amounts of floor space attributed to each department. In total, we model three variables: total sales \( S_{jt} \), departments’ sales shares \( DSS_{mit} \), and departments’ relative floor space sizes \( SS_{mit} \). In the remainder of this section, we present and discuss these models, which we use to explain the variables.

4.3.1 Overall Sales

To evaluate the impact of outlet location on overall sales, we use a Tobit model. Overall sales per zip code are bounded by zero and skewed to the right. If we analyze the amount of sales generated per zip code, we should also account for non-
negativity and the large number of zero observations. Tobin (1958) developed a model to explain this type of variable, taking the following form:

\[ S_{jt}^* = \alpha_0 + \sum_{q=1}^{Q} \alpha_q X_{jq} + \sum_{j=1}^{L} \beta_j Z_{jtl} + \varepsilon_{jt}, \]  

(4.2a)

and

\[ S_{jt} = \begin{cases} S_{jt}^* & \text{if } S_{jt}^* > 0 \\ 0 & \text{if } S_{jt}^* \leq 0, \end{cases} \]

(4.2b)

where \( S_{jt}^* \) is a latent variable measuring the amount of sales generated at a particular location, which can be negative, positive, or zero. However, if the (unobserved) sales in a particular zip code, as predicted by Equation 4.2a, are negative, \( S_{jt} \), the observed sales level will be zero, as formalized by Equation 4.2b.

The set of explanatory variables includes variables observed at the store (\( X \)) and zip code (\( Z \)) levels. The store-specific explanatory variables are measured such that they refer to the nearest store.

We extend the regular Tobit model to include spatially autocorrelated error terms, because zip codes in close proximity often share unobservable characteristics (e.g., history, resources, infrastructure), and consumer spending levels in neighboring zip codes cannot be considered fully independent. If the observations \( (j = 1, \ldots, J) \) are stacked in a vector for each cross-section of zip codes at time \( t \), we can account for spatial error autocorrelation by

\[ \varepsilon_i = \delta W \varepsilon_i + \xi_i, \]

(4.3)

where \( E(\xi_i) = 0 \), \( Var(\xi_i) = \sigma^2 I_J \), and \( W \) is a row-standardized first-order contiguity weight matrix of size \( (J \times J) \) that describes the spatial arrangement of zip codes.
4.3.2 Department Sales Shares

We use an attraction model to explain a department’s sales share at a particular location, based on its relative size compared with the nearest store, other store attributes, and competitor and consumer characteristics observed at the zip code level. Attraction models are useful tools to analyze competitive interactions (Carpenter et al. 1988; Cooper and Nakanishi 1988; Nakanishi and Cooper 1982), because of their logical consistency; that is, market shares sum to unity, and the market shares of individual brands are between 0 and 1. Campo and Gijsbrechts (2004) and Campo et al. (2000) have used the attraction model for purposes similar to ours, but rather than explaining sales shares at the department level, they explain product category sales shares. In our setting, the attraction model takes the following form:

\[
DSS_{mjt} = \frac{A_{mjt}}{\sum_{c=1}^{M} A_{cjt}},
\]

(4.4a)

where

\[
A_{mjt} = \exp(\beta_{m} + \varepsilon_{mjt}) SS_{mjt} \prod_{k=1}^{K} \exp(\gamma_{km} X_{ik}) \prod_{n=1}^{N} \exp(\lambda_{nm} Z_{jin}),
\]

(4.4b)

and \( A_{mjt} \) is the attraction of department \( m \) in zip code \( j \) at time \( t \), \( SS_{mjt} \) is the fraction of store space devoted to department \( m \) in the store closest to zip code \( j \) at time \( t \). The set of explanatory variables includes variables observed at the store (\( X \)) and at the zip code level (\( Z \)). The store-specific explanatory variables (\( X \)) will be measured such that they refer to the nearest store.

4.3.3 Department Sizes

As we noted in the beginning of Section 4.3, there is a potential endogeneity problem with the models. In practice, a retailer may decide to allocate more floor space to departments that are selling well in a particular store (Van Dijk et al. 2004;
Van Nierop, Fok, and Franses (2008). In this case, the store manager might use previous department sales (shares) to determine the optimal assortment composition; that is, the (relative) amount of space attributed to a department is a function of its past performance. If this endogeneity is ignored, we would likely overestimate the impact of changes in a department’s (relative) size on its share of sales. In technical terms, the explanatory variables $SS_{mjt}$ in Equation 4.4b are not uncorrelated with the error term of this equation. To correct for this point, department sizes should be considered endogenous. We therefore specify a model for assortment composition, in which department sizes are considered a function of store, competitor, and (aggregated) consumer characteristics. We thus again adopt an attraction model specification, because relative department sizes also satisfy the logical consistency requirements of this model type. Hence, we model $SS_{mit}$ using:

$$SS_{mit} = \frac{Att_{mit}}{\sum_{c=1}^{M} Att_{cit}}$$  \hspace{1cm} (4.5a)

$$Att_{mit} = \exp(\rho_m + \nu_{mit}) \prod_{g=1}^{G} \exp(\delta_{gm} X_{itg}) \prod_{r=1}^{R} \exp(\phi_{rm} \bar{Z}_{itr})$$  \hspace{1cm} (4.5b)

in which the $\bar{Z}_{itr}$ variables are the averages for each zip code–level variable for all zip codes for which store $i$ is the nearest store. The set of store-specific explanatory variables now includes a variable measuring the one-period lag of department $m$’s share of sales.

We include explanatory variables that capture a store’s profile, which is defined as characteristics of the store, consumers, and competitors. Van Dijk et al. (2004) find that retailers are more likely to allocate similar amounts of shelf space to

---

1 Note that the notation for the relative department sizes in Equation 4.5 differs from that in Equation 4.4, because we include a superscript $i$ in Equation 4.5 instead of the subscript $j$ in Equation 4.4. The $SS_{mit}$ refers to the relative size of department $m$ in a particular store $i$ at time $t$, whereas $SS_{mjt}$ measures the relative size of department $m$ in the store closest to zip code $j$ at time $t$.
brands for stores with comparable profiles. We include store profiles as explanatory variables rather than, as Van Dijk and colleagues (2004) do, an operationalization of the spatial weights matrix in which the weights are distances derived from a multidimensional scaling analysis. We believe that our specification offers richer insights, because we can assess the impact of each location variable separately rather, than on just the two dimensions obtained through a principal components analysis.

Following Bronnenberg and Mahajan (2001), we also account for (unobserved) spatial dependencies in relative department sizes across stores. In particular, we assume the unexplained part of department $m$’s relative size in store $i$ to be a function of those of neighboring stores. The rationale behind this assumption is that stores in close proximity should share unobservable characteristics that may lead to similar fractions of floor space devoted to each department.

4.4 Attraction Model Estimation

Estimating the parameters of an attraction model is not straightforward; we must transform the dependent and explanatory variables to obtain a model that is linear in parameters and satisfies logical consistency conditions. Fok, Franses, and Paap (2002) show that this model can be achieved by considering Equations 4.4b and 4.5b as the $m^{th}$ equations in a set of $M$ equations. Because the sales shares of all departments in a particular zip code (and all department shares of total floor space) by definition sum to 1, dependencies across equations exist, so we do not have a full rank system. We linearize the system in Equations 4.4 and 4.5 to an equivalent system of $M - 1$ equations by arbitrarily selecting a base brand $M^*$ and taking the ratio between $DSS_{mji}$ and the share of this brand:
If we then take the natural logarithm of both sides, we obtain a system of $M - 1$ equations that is linear in the parameters. For notational convenience, we define the log-transforms $\log m_{jt} = \log(S_{mjt})$, log-ratios $\tilde{y}_{mjt} = \log(DSS_{mjt}/DSS_{M^*,jt})$, and the following differences: $\tilde{\beta}_{1m} = \beta_{1m} - \beta_{1M^*}$, $\tilde{y}_{km} = \gamma_{km} - \gamma_{km^*}$, $\tilde{\lambda}_{km} = \lambda_{km} - \lambda_{km^*}$, and $\eta_{mjt} = \epsilon_{mjt} - \epsilon_{M^*,jt}$. These tactics simplify Equation 4.6 to

$$
\tilde{y}_{mjt} = \tilde{\beta}_{1m} + \beta_{2m}s_{mjt} - \beta_{2M^*}s_{M^*,jt} + \sum_{k=1}^{K}\tilde{y}_{km}X_{jkt} + \sum_{n=1}^{N}\tilde{\lambda}_{nm}Z_{jtn} + \eta_{mjt},
$$

for $m = 1, \ldots, M - 1$. Consequently, we can only estimate the parameters $\tilde{\beta}_{1m}, \beta_{2m}, \beta_{2M^*}, \tilde{y}_{km}$, and $\tilde{\lambda}_{km}$, not (all) the model parameters in Equations 4.4 and 4.5. Yet the identification of these reduced-form parameters is sufficient to calculate elasticities (Cooper and Nakanishi 1988; Fok, Franses, and Paap 2002).^{2}

We assume the $M - 1$ equations to be correlated, because a zip code with high unobserved variables for the sales percentage of one department probably also has them for other departments. The transformed disturbances $\eta = (\eta_{1}\ldots\eta_{M-1})$ therefore follow a normal distribution with mean zero and covariance matrix $\tilde{\Sigma} = L\Sigma L'$. Also, $L = (I_{M-1}, i_{M-1})$, in which $I_{M-1}$ is a $(M - 1)$-dimensional identity matrix, and $i_{M-1}$ is a $(M - 1)$-dimensional vector; therefore, only $\frac{1}{2}M(M - 1)$ parameters of the original covariance matrix $\Sigma$ of the error terms $\epsilon = (\epsilon_{1}\ldots\epsilon_{M})$ can be identified. Each equation contains a unique set of explanatory variables, so

---

^{2} The coefficient $\beta_{2M^*}$ is equal across the $(M - 1)$ equations. This restriction is taken into account when the parameters are estimated.
estimating each equation separately by ordinary least squares would lead to inefficient parameter estimates. We use a feasible generalized least squares (GLS) procedure, better known as the seemingly unrelated regressions (SUR) model, to estimate the model parameters. The reduced-form parameters can calculate elasticities (Fok, Franses, and Paap 2002), and the elasticity of department \( m \) ’s relative size in store \( i \) on its percentage sales in zip code \( j \) equals:

\[
\frac{\delta DSS_{mit}}{\delta SS_{mit}} \frac{SS_{mit}}{DSS_{mit}} = \left(1 - DSS_{mit}\right) \beta_{2m},
\]

(4.8)

According to Equation 4.8, elasticity converges to 0 if the department’s share of sales goes to 1. If a department’s sales share is a increasing function of its relative size in the nearest store, that is, \( \beta_{2m} > 0 \), then the elasticity will go to 0 if \( SS_{mit} \) goes to infinity. We can increase a particular department’s sales share by enlarging the floor space attributed to that department, but for higher floor space levels, the impact of space on department sales quickly levels off. This relationship, which exhibits decreasing returns to scale, is consistent with prior store space allocation literature (Desmet and Renaudin 1998).

We extend the reduced-form model in Equation 4.7 with spatial error autocorrelation to account for spatial dependencies in (the log of) each department’s relative sales shares across zip codes. The spatial error model posits that (the log of) a department’s relative sales share in a particular zip code depends on unobservable characteristics, correlated across space, as a result of which the error terms follow a spatial first-order autoregressive process that generates the error terms,

\[
\eta_{mt} = \kappa_{m} W_{t} \eta_{mt} + \xi_{mt},
\]

(4.9)

where \( \eta_{mt} \) is a \( (J_{t} \times 1) \) vector of spatially correlated error terms for every zip code \( (j = 1, \ldots, J_{t}) \), and \( \xi_{mt} \) is a vector of error terms which are not correlated over space. Furthermore, \( \kappa_{m} \) is the spatial autocorrelation coefficient, which can differ between departments, and \( W_{t} \) denotes a \( (J_{t} \times J_{t}) \) non-negative matrix with zeros
on the diagonal that describes the spatial arrangement of zip codes. Note that the spatial weights matrix \( W_t \) differs for each year, as indicated by the subscript \( t \), because a department’s sales shares in a particular zip code are observed only if \( S_{jt} > 0 \). Therefore, the number of zip codes \( J_t \) for which shares must be explained differs for each time period.

### 4.5 Prediction

An important element of the proposed modeling approach involves predicting the sales impacts of future changes in a store’s retail environment and location and assortment changes, as well as new store openings. Prediction from the models is not straightforward though. Fok, Franses, and Paap (2002) show, for example, that it is impossible to calculate the expected values of department sales shares analytically; they can be obtained only through simulations. The market shares averaged over a large number of iterations then provide the basis for calculating expectations.

We therefore randomly draw the \((J \times 1)\) vector \( \xi_{mt}^d \) a certain number of \( D \) times, \( d = 1, \ldots, D \), from the estimated covariance matrix \( \tilde{\Sigma} \) to obtain \( \eta_{mt}^d \) by

\[
\eta_{mt}^d = \left( I_{J_t} - \kappa_m W_t \right)^{-1} \xi_{mt}.
\]

We use these disturbances and the parameters of the reduced-form model to predict relative market shares \( dss_{mjt} = DSS_{mjt} / DSS_{M^*,jt} \) in zip code \( j \) for department \( m \):

\[
dss_{mjt}^d = \exp\left( \beta_{lm} + \eta_{mjt}^d \right) SS_{mjt}^{\beta_{lm}} \prod_{k=1}^{K} \exp\left( \tilde{\gamma}_{km} X_{jkt} \right) \prod_{n=1}^{N} \exp\left( \tilde{\lambda}_{mn} Z_{jtn} \right). \tag{4.10}
\]

Because \( dss_{M^*,jt}^d = 1 \), we can compute each department’s share in the total amount of sales generated in a particular zip code by using the following equation:
Provided that we use a sufficiently large number of draws (\(D\)), we can finally approximate the expected values of the department shares by taking the average of the sales shares over all draws.

If we use the Tobit model in Equation 4.2 to predict overall sales at the zip code level, we also

\[
E\left(S_{jt} \mid U_{jt}\right) = P\left(S_{jt} = 0 \mid U_{jt}\right) E\left(S_{jt} \mid S_{jt} = 0, U_{jt}\right) + P\left(S_{jt} > 0 \mid U_{jt}\right) E\left(S_{jt} \mid S_{jt} > 0, U_{jt}\right),
\]

(4.12)

where \(U_{jt} = [X_{jt}, Z_{jt}]\), \(\theta = [\alpha_0, \alpha, \beta']\), \(\alpha\) is a \((Q \times 1)\) vector of \(\alpha_q\) s, and \(\beta\) is a \((L \times 1)\) vector of \(\beta_j\) s. Because \(E\left(S_{jt} \mid S_{jt} = 0, U_{jt}\right)\) equals 0 if \(S_{jt}\) is censored from below, \(P\left(S_{jt} > 0 \mid U_{jt}\right) = \Phi\left(\frac{U_{jt}\theta}{\sigma}\right)\), and

\[
E\left(S_{jt} \mid S_{jt} > 0, U_{jt}\right) = U_{jt}\theta + \sigma \frac{\varphi\left(U_{jt}\theta/\sigma\right)}{\Phi\left(U_{jt}\theta/\sigma\right)}. 
\]

In turn, we obtain the following expression for expected sales in zip code \(j\) at time \(t\):

\[
E\left(S_{jt} \mid U_{jt}\right) = U_{jt}\theta \Phi\left(U_{jt}\theta/\sigma\right) + \sigma \varphi\left(U_{jt}\theta/\sigma\right),
\]

(4.13)

where \(\Phi(z)\) and \(\varphi(z)\) denote the cumulative density function (cdf) and probability density function (pdf) of the standard normal distribution, respectively.

4.6 Data

In addition to the store-level variables of the attraction model, we include the number of households living in a particular zip code as an additional covariate in the sales model.

The data set analyzed in this chapter also contains information about 30 stores belonging to a Dutch clothing chain. The chain’s positioning is targeted at middle-
class families, as reflected in the stores’ average price levels and medium-quality assortments for men, women, and children. The chain uses a loyalty program to strengthen its relationship with regular customers. Participants receive a 5 percent cash reward on every purchase, credited to their loyalty cards, which can be spent freely two times a year. Although all stores offer clothes for men, women, and children, the relative amount of floor space devoted to each department differs for each store. Store size ranges from approximately 500 m² to 2530 m², of which an average of 44 and 21 percent contains women’s and children’s clothes, respectively.

From the chain’s customer database, we collected yearly data on department-level sales for all zip codes in the Netherlands in five successive years (2002–2006); we use the first four years for estimation and the last year for validation. These data are supplemented with survey data on the retail environment in which each store operates and commercially available geodemographic information. We identify the number of competitors for each store using information obtained from a survey among store managers. Other store attributes include store size (in 10,000 m²), the relative sizes of the various departments, the number of months a store is open in a particular year, and store age (1932 = 0). We do not include other marketing mix variables, because store managers must adhere to the marketing activities dictated by the head office.

We finally use a wide variety of socio-demographic variables observed at the zip code level to evaluate the impact of consumer characteristics on overall and department-level sales. These variables can influence store performance in several ways. First, variables such as the number of households living in a particular area and their socio-economic status determine local buying power and affect overall spending levels. Second, other variables may drive need patterns and differentially affect the sales level of each department. An example would be if many households with children create a higher demand for children’s clothes in a particular region. Third, consumer variables may drive store patronage, because channels (or
individual stores) differ in their attractiveness to certain consumer groups (Inman, Shankar, and Ferraro 2004). This effect is partly captured through the inclusion of variables that measure the amount of floor space allocated to each department in our model as a means to explain overall sales levels. The assortment composition determines a store’s attractiveness to certain consumer groups, which in turn affects its sales level.
Table 4.2: Parameter estimates of attraction model explaining relative department sizes

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>Dependent Variable</th>
<th>Women’s Department Size</th>
<th>Coeff</th>
<th>t-value</th>
<th>Men’s Department Sales Share</th>
<th>Coeff</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td>3.854</td>
<td>2.91***</td>
<td>2.165</td>
<td>2.20**</td>
<td></td>
</tr>
<tr>
<td>Store characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size (in 10,000 m²)</td>
<td></td>
<td>-7.418</td>
<td>-4.29***</td>
<td>-2.196</td>
<td>-1.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged sales female assortment</td>
<td></td>
<td>0.966</td>
<td>3.85***</td>
<td>0.963</td>
<td>5.43***</td>
<td></td>
<td>0.963</td>
</tr>
<tr>
<td>Lagged sales male assortment</td>
<td></td>
<td>-0.750</td>
<td>-1.48</td>
<td>-0.750</td>
<td>-1.48</td>
<td></td>
<td>-0.750</td>
</tr>
<tr>
<td>Proportion of the year the store is open (in months)</td>
<td></td>
<td>-0.768</td>
<td>-1.81</td>
<td>-0.501</td>
<td>-1.52</td>
<td></td>
<td>-0.501</td>
</tr>
<tr>
<td>Year of establishment (/100)</td>
<td></td>
<td>-0.737</td>
<td>-3.27***</td>
<td>-0.355</td>
<td>-2.23**</td>
<td></td>
<td>-0.355</td>
</tr>
<tr>
<td>Competitor characteristics</td>
<td></td>
<td>No of competitors female ass. (/100)</td>
<td>-0.287</td>
<td>-1.14</td>
<td>0.570</td>
<td>1.65</td>
<td>0.570</td>
</tr>
<tr>
<td>Consumer characteristics</td>
<td></td>
<td>% households with high SES</td>
<td>-2.055</td>
<td>-1.41</td>
<td>2.959</td>
<td>2.68***</td>
<td></td>
</tr>
<tr>
<td>% households with low SES</td>
<td></td>
<td>-0.050</td>
<td>-0.03</td>
<td>3.577</td>
<td>2.99***</td>
<td></td>
<td>3.577</td>
</tr>
<tr>
<td>% of foreigners</td>
<td></td>
<td>2.711</td>
<td>2.74***</td>
<td>1.244</td>
<td>1.56</td>
<td></td>
<td>1.244</td>
</tr>
<tr>
<td>% couples</td>
<td></td>
<td>-5.245</td>
<td>-2.92***</td>
<td>-6.222</td>
<td>-4.37***</td>
<td></td>
<td>-6.222</td>
</tr>
<tr>
<td>% households with children</td>
<td></td>
<td>-0.937</td>
<td>-0.71</td>
<td>0.171</td>
<td>0.17</td>
<td></td>
<td>0.171</td>
</tr>
<tr>
<td>D2004</td>
<td></td>
<td>-0.021</td>
<td>-0.50</td>
<td>0.006</td>
<td>0.14</td>
<td></td>
<td>0.006</td>
</tr>
<tr>
<td>D2005</td>
<td></td>
<td>-0.036</td>
<td>-0.79</td>
<td>-0.022</td>
<td>-0.46</td>
<td></td>
<td>-0.022</td>
</tr>
<tr>
<td>Spatial autocorrelation coeff (λ)</td>
<td></td>
<td>-0.522</td>
<td>0.00</td>
<td>0.054</td>
<td>0.00</td>
<td></td>
<td>0.054</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td></td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td>84</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *** p < 0.01 ** p < 0.05
<table>
<thead>
<tr>
<th>Table 4.3: Parameter estimates of models explaining department sales shares and total sales</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Explanatory Variable</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
</tr>
<tr>
<td><strong>Store characteristics</strong></td>
</tr>
<tr>
<td>m² female assortment</td>
</tr>
<tr>
<td>m² male assortment</td>
</tr>
<tr>
<td>m² children’s assortment</td>
</tr>
<tr>
<td>Proportion of the year the store is open (in months)</td>
</tr>
<tr>
<td>Year of establishment (/100)</td>
</tr>
<tr>
<td><strong>Competitor characteristics</strong></td>
</tr>
<tr>
<td>No of competitors female ass. (/100)</td>
</tr>
<tr>
<td>Total no of competitors (/100)</td>
</tr>
<tr>
<td><strong>Consumer characteristics</strong></td>
</tr>
<tr>
<td>Distance to the store (in miles)</td>
</tr>
<tr>
<td>Distance to next-nearest store (in miles)</td>
</tr>
<tr>
<td>Number of households (/1,000)</td>
</tr>
<tr>
<td>Average household size</td>
</tr>
<tr>
<td>% households with high SES</td>
</tr>
<tr>
<td>% households with low SES</td>
</tr>
<tr>
<td>% of foreigners</td>
</tr>
<tr>
<td>% households with children</td>
</tr>
<tr>
<td>D2004</td>
</tr>
<tr>
<td>D2005</td>
</tr>
<tr>
<td>Spatial autocorrelation coeff (λ)</td>
</tr>
<tr>
<td>R²</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

Notes: *** p < 0.01 ** p < 0.05
4.7 Drivers of Department Sales: Estimation Results

4.7.1 Total Sales

The parameter estimates reported in the last column of Table 4.3 indicate that the overall sales level at a particular zip code negatively depends on the distance to the nearest store in the focal chain. This finding is consistent with prior literature on spatial interaction models, which assumes that the probability of visiting a particular store is inversely related to the distance to the store. The theory of the allocation of time advocated by Becker (1965) also predicts that consumers perceive disutility from traveling, due to transportation and opportunity costs, and therefore are more likely to visit stores closer to their residence (Bawa and Ghosh 1999; Bhatnagar and Ratchford 2004).

We find a significant positive effect of the number of households on sales. Regions with large populations constitute potentially large markets, and retail outlets can generate more sales from these regions (Kumar and Karande 2000; Reinartz and Kumar 1999). Zhu and colleagues (2009) go as far as to conclude that population size is the single most important determinant of market structure. They find that in the retail discount industry, markets with no stores have significantly smaller populations than markets with stores. This finding implies that retailers use population size as a proxy for potential sales and employ such predictions to make their entry decisions. Moreover, households with children and couples spend significantly more on clothes than do single-person households at this retailer, likely because larger families have more diverse needs and buy a wider variety of products, which on average produces higher sales for these consumer groups (Bawa and Ghosh 1999).

We also find a positive and significant effect of the number of competitors on the chain’s overall sales level, which means that more rival stores enhance target store performance. This finding contradicts recent results from Zhu and Singh
(2009), who indicate that competition typically exerts a strong negative effect on store performance. Although previous studies show that the effect of competition between stores is lower if the market is large enough to support multiple stores, stores located farther away and/or those with different retail formats (Ailawadi et al. 2009; Gielens et al. 2008; Zhu, Singh, and Manuszak 2009) could produce a positive effect of competition. Nelson’s theory of cumulative attraction assumes that stores in close proximity earn more business than those located far apart, because consumers visiting multiple stores have a lower risk of product unavailability (González-Benito and González-Benito 2005). Because shopping for clothes is sometimes regarded as a recreational activity, individual stores also may benefit from the presence of competitors, who offer the promise of comparison shopping (Dholakia 1999).

4.7.2 Department Shares

The second and third columns of Table 4.3 contain the parameter estimates for the models that explain sales shares in the women’s and men’s departments. We estimated these models using a reduced-form specification in which the children’s department was the reference category, so a positive (negative) coefficient for a particular variable means that it affects the sales share of this department more (less) strongly than does the share of the children’s department.

The results show that more floor space allocated to a particular department increases the sales share for this particular department. This finding confirms the results of, among others, Campo et al. (2000), Desmet and Renaudin (1998), and Van Nierop, Fok, and Franses (2008). A possible explanation for this effect indicates that products in departments with a large share of the store’s total floor space receive more attention from consumers and are more likely to appear in their shopping baskets (Desmet and Renaudin 1998). If more floor space is allocated to a particular department, the retailer also can display more items, giving consumers
more products from which to choose so that they are more likely to find what they want and increase sales in this particular department (Hoch et al. 1995). Similarly, more space allows the retailer to hold extra inventory and lower the risk of out-of-stocks, which may have a positive effect on sales (Desmet and Renaudin 1998).

The sales shares of the men’s and women’s department are positively affected by the number of competitors in each department. If we hold total sales constant, the men’s and women’s departments thus benefit more from the presence of competitors than does the children’s department and obtain a larger share of overall sales. Travel distance also positively affects these sales shares; men’s and women’s departments achieve higher sales shares when consumers live farther from the store. The combination of these findings suggests that men’s and women’s departments benefit most from the spatial concentration of apparel stores.
Figure 4.1: Predicted sales levels for each zip code in the Netherlands in the year 2006.
The left panel shows the sales distribution if there are no new stores, whereas the right panel depicts the predictions for total chain sales in a situation with two new stores.
4.7.3 Department Sizes

In Table 4.2, we present the parameter estimates for the models that explain the relative sizes of the men’s and women’s departments. We include this model to account and control for potential endogeneity in the store space allocation decision, which results when store managers allocate larger (smaller) amounts of store space to better (worse) performing departments. We find evidence of such effects, as indicated by the positive coefficients that measure the effects of (lagged) department sales on the amount of floor space allocated to this department. If departments have high (past) sales levels, more space gets allocated to them. We thus corroborate the findings of Van Dijk et al. (2004) and Van Nierop, Fok, and Franses (2008).

In addition, larger stores appear more likely to have a (larger) children’s department. The year of establishment variable indicates a negative relationship with the relative sizes of the men’s and women’s departments; newer stores have relatively larger children’s departments. We also note that though the chain’s format generally should appeal to middle-class families, store managers are more likely to enlarge the men’s department if the proportion of households with low and high socio-economic status greater large. The focal stores thus could be more attractive to (single) men in these consumer groups.

4.8 Potential Application: Store Location Evaluation

We now know not only which location variables drive the total amount of sales generated in a zip code but also how the performance of each individual department is affected by each variable. To examine whether the proposed model can predict sales correctly for new stores, we use the holdout sample of two newly opened stores in 2006. These two new stores have similar characteristics to the other stores and appeared in the midwestern part of the Netherlands (in Figure 4.1, numbers 1 and 2). Adding these new stores to the data set implies that the variables measuring
the characteristics of and travel distances to the nearest store should change for a substantial number of the zip codes located in the western part of the country.

To see how the distribution of a department’s sales changes after the two new stores open, we use the coefficients reported in the last column of Table 4.3 to predict the total sales for each four-digit zip code in the Netherlands for a hypothetical situation in which the new stores are not present in 2006 but the values of the variables measuring consumer characteristics represent those observed in 2006. The predicted spatial distribution of sales appears in Figure 4.1a. We also update the set of explanatory variables so that the variables measuring travel distances to and characteristics of the nearest store include the new stores, with the results in Figure 4.1b for the predicted sales distribution. To evaluate and compare the predicted amount of sales with the observed sales figures, we subtract the amount of sales generated in a zip code before from that after the opening of the two new stores. The next step is to determine the geographical extent of the trade area for each store, then sum the sales for all zip codes within the store’s trade area. To determine the size of a store’s trade area, we use the same trade area perimeters as in Chapter 3, that is, the maximum travel distance to the store for the first zip codes responsible for 85% of total sales. These distances are 13.36 and 14.17 for stores 1 and 2, respectively. If we sum the sales levels for all zip codes that belong to the trade area for each store, we can compare the realized sales figures with the predicted values. As we show in Figure 4.2, the predicted sales figures are very close to the realized values. The predicted sales for store 1 are €1,109,984, very close to the predicted sales (€988,320), and for store 2, actual sales equal €250,406, very close to the predicted value (€154,943). Therefore, the model predictions for total sales approximate the observed values well, so the model is useful for store location evaluations.
4.9 Scenario Analysis: Relative Department Size

Our modeling approach also enables us to determine what happens to total sales if we change the sizes of store departments. We use the coefficients in the rightmost column of Table 4.3 to predict the overall sales level for each zip code belonging to the trade area of store 1 but change the variables that measure department sizes, setting all other predictor variables to their average values. To quantify the impact on total store sales, we change the size of each department at increments of 5 m². In Figure 4.3, we depict the predicted sales levels for several combinations of individual department sizes, assuming total floor space does not increase. An increase in the sizes of the children’s and women’s departments enhances total store sales. Specifically, increasing the size of the women’s department by 1 m² has an effect similar in size (€15,792) to a similar enlargement of the children’s department (€15,825). Increasing the size of the men’s department negatively affects store sales though. This particular store therefore might increase its potential sales by enlarging the proportion of floor space it devotes to the children’s and women’s departments, at the expense of the men’s department.

Figure 4.2: Observed and predicted sales figures for two stores opened in 2006.
4.10 Conclusions and Discussion

Travel distance to the store and assortment variety are the two of the most important factors that consumers consider when deciding where to shop (Briesch, Chintagunta, and Fox 2009). Retailers use these elements of the marketing mix to differentiate themselves from competitors, so they determine the structure of the local market and potential profit levels. In this sense, location choice and assortment composition are critical elements of the marketing mix that must be determined in combination. We therefore propose a model for store location evaluations that acknowledges the moderating effect of location characteristics on the optimal assortment composition for each store.

The model we propose in this chapter contributes to prior literature in several ways. We extend micromarketing literature, in that we evaluate the performance implications of changes in new store assortments, a consideration never adopted previously (Campo et al. 2000). We also allow for more heterogeneity in consumer characteristics than currently available models offer, in that we use consumer data observed at the zip code level rather than aggregated socio-demographic profiles for

Figure 4.3: Response of total store sales to a change in the size of the children’s and men’s department.
each store (Kumar and Karande 2000). Moreover, we account for spatially correlated error terms that may result from unobserved imitating behavior by consumers (Choi, Hui, and Bell 2010), retailers (Bronnenberg and Mahajan 2001), or other variables that cause spatial dependence in department sales levels across zip codes.

We test the proposed model using data from a Dutch clothing chain that operates 30 stores in various locations, each offering a retail assortment to middle-class families. The empirical study confirms previous findings from Campo et al. (2000): Location variables affect each department’s sales shares differently. Travel distance to the store, for example, affects the sales shares of the men’s and women’s department more strongly than that of the children’s department, whereas in areas with more families with children, the sales share of the children’s department is greater. Total sales levels are higher in areas closer to the store, where there is intense competition and greater market potential (i.e., number of households). We further find evidence that retailers decide about the amount of floor space devoted to each department, based on each department’s past performance. The size of each department in a particular year positively depends on its sales level in the previous year, so department sizes are endogenous, and we would likely overestimate the sales impact of changes in the amount of floor space allocated to each department if we were to ignore this reverse effect.

Not only does this study increase our understanding of the performance implications of tailoring assortments to local store environments, but it also has some limitations that should be addressed in further research. First, we consider only one chain of stores; the findings are therefore peculiar to the positioning of this particular chain and difficult to generalize. Second, we do not really optimize the overall performance of the chain. To do so, we would need data about the (average) unit (gross) margins for each individual store department and an estimate of the costs associated with adding new stock to the assortment of each department. Third,
our modeling approach does not take into account the potential endogeneity of market structure (Zhu and Singh 2009), which would imply that we cannot assume the number of competitors, their locations, and their assortments are given. Previous research has shown that retailers note the (anticipated) location choices of their competitors when they choose locations; if we wanted to identify the optimal location and assortment for each store, we should investigate potential (future) reactions of competitors and solve location and assortment decisions for several retailers simultaneously rather than sequentially. Fourth, we consider location and assortment decisions—just two elements of the marketing mix. Finally, we consider a model with limited competition assuming that a department’s attractiveness only depends on its own explanatory variables and not on those of other departments. Previous research shows the existence of cross-demand effects at the product category level (Leeflang and Parreño Selva 2010). Future research is needed to investigate whether similar effects can be observed at the level of store departments. Additional research should address local marketing strategies for other elements, such as prices (Hoch et al. 1995) and promotions (Gijsbrechts, Campo, and Goossens 2003).

Despite these limitations, we believe that the proposed model can be very valuable to retailers that want to open new chain stores and tailor their assortments to local conditions. As we have shown in our empirical study, this model effectively predicts potential sales by new store locations and the sales impacts of changes in the assortment composition, which makes it a useful tool to support these decisions.