Chapter 4
Existing Models and Store Profile Data

4.1 Introduction
In the previous chapter we have seen that the models practitioners use for local marketing focus on sales potential. This potential is used as a basis for the composition of the marketing mix. Models that explain marketing mix effects have, as far as we know, not yet been implemented by practitioners. These models have been developed by marketing scientists (see also Bucklin and Gupta, 1999). In this chapter, we study existing models and the available data on store profile variables. In the next section we describe some models that are used in Dutch practice. Section 4.3 studies models developed by marketing scientists that estimate store-specific marketing mix effects. We discuss available store profile data in the Netherlands in Section 4.4.

4.2 Models in practice
We illustrate the modeling approach in practice by explaining models and services that relate store profiles to store sales. We focus on the products offered by ACNielsen in the Netherlands. These products are FamilyTrack, Local Marketing and Assortman.

FamilyTrack is a service that provides descriptive marketing data for different clusters of supermarkets. These clusters are based on social class and family life cycle. The marketing data refer to variables such as sales, distribution and promotions. FamilyTrack allows the user to compare clusters of supermarkets and use these data to customize the marketing mix at this level. FamilyTrack describes rather than explains the market. It does not use store profile data to predict sales or marketing mix effects for individual stores.

Local Marketing is a model that uses store profiles to explain sales at the store level from a regression model. This provides insight into which store profile variables are important and what sales levels can be expected for a particular store given its profile (potential sales). A comparison of the potential and actual sales
Chapter 4

provides insight into how the store performs relative to other stores. This comparison can be used to give an advice on the marketing mix. This service however, does not explain marketing mix effects from store profiles.

Assortman is a service that optimizes the assortment in a store. In most cases the focus is on the optimal number of SKUs (Stock Keeping Units). The model can be implemented at different aggregation levels including categories, subcategories, and segments.

Assortman consists of two steps. First, SKU elasticities are determined. The elasticities are assumed to be constant across stores and are inferred using cross-sectional data. The model uses store-profile variables to control for differences between stores. Second, the assortment is optimized using the elasticity estimates. The optimum is compared to the allocation that can be expected based on the store profiles: the expected allocation is estimated from a different model as data on the actual allocation are not available.

Importantly, the focus of Assortman is on the assortment and not on local marketing. Assortman does not account for store profile dependent allocation elasticities.

4.3 Models in academics

There are four studies in the marketing literature that use store profiles to determine store-specific marketing mix effects. The first three studies, by Hoch et al. (1995), Montgomery (1997), and Mulhern et al. (1998) focus on price elasticities. Hoch et al. (1995) and Mulhern et al. (1998) explain price elasticities from store profile variables. Montgomery (1997) explains all parameters including price elasticities. In addition, he shows how to optimize the marketing mix. Campo et al. (2000) consider the relationship between store profiles and area allocation across categories. They show how the store area should be allocated to enhance store sales.

Hoch et al. (1995)

Hoch et al. (1995) use a two-step procedure to model store-specific category level price elasticities as a function of store profiles. In the first step, they estimate store-specific price elasticities per product category. In the second step, they relate these elasticities to consumer and competitor characteristics.

Hoch et al. (1995) determine category level price elasticities as aggregated brand level estimates. They use brand level estimates to avoid the extreme
Existing models and store profile data

assumptions and possible aggregation bias in an aggregated analysis. For each category, they define a log-linear demand system:

\[
\ln SALES_{kt} = \alpha + \tau_k \cdot t_j + (N + A_k) \ln PRICE_{kt} + \Phi \ln SALES_{kt-1} + \Psi \cdot DEAL_{kt} + \Xi \cdot FEAT_{kt} + \epsilon_{kt} \tag{4.1}
\]

with

- \(SALES_{kt}\) = \(J \times 1\) vector of brand sales in store \(k\) in week \(t\). Elements of the vector are the \(SALES_{jkt}\), the sales of item \(j\) in store \(k\) in week \(t\), \(j = 1, \ldots, J, k = 1, \ldots, K, t = 1, \ldots, T\),
- \(PRICE_{kt}\) = \(J \times 1\) vector of prices in store \(k\) in week \(t\). Elements of the vector are the \(PRICE_{jkt}\), the price of item \(j\) in store \(k\) in week \(t\),
- \(DEAL_{kt}\) = \(J \times 1\) vector of dummy variables that indicate temporary price reductions or in-store coupons in store \(k\) in week \(t\). Elements of the vector are the \(DEAL_{jkt}\), the indicator for item \(j\) in store \(k\) in week \(t\),
- \(FEAT_{kt}\) = \(J \times 1\) vector of dummy variables that indicate whether there is a feature. Elements of the vector are the \(FEAT_{jkt}\), the indicator for item \(j\) in store \(k\) in week \(t\),
- \(\alpha\) = \(J \times 1\) vector with item specific intercepts. Elements of the vector are the \(\alpha_j\), the intercept for item \(j\),
- \(\tau_{k \cdot t_j}\) = \(J \times 1\) vector with all elements equal to the store-specific intercept, \(\tau_k\). \(t_j\) is a \(J \times 1\) vector of ones,
- \(N + A_k\) = \(J \times J\) matrix of own- and cross- price elasticities,
- \(\Phi\) = \(\phi I\) a \(J \times J\) diagonal matrix of lagged price effects, all effects are equal to \(\phi_1^{1}\),
- \(\Psi\) = \(J \times J\) diagonal matrix with on the diagonal item specific own-deal effects \(\delta_{jj}\),
- \(\Xi\) = \(J \times J\) diagonal matrix with on the diagonal item specific feature effects \(\psi_{jj}\),
- \(\epsilon_{kt}\) = \(J \times 1\) vector of random disturbances in store \(k\) in week \(t\). Elements of this vector are the \(\epsilon_{jkt}\), the disturbance for item \(j\) in store \(k\) in week \(t\).

\(^1\) The lagged sales variable is included to accommodate for serial correlation caused by forward buying or stockpiling behavior.
Chapter 4

The demand system allows for a full pattern of cross elasticities captured by the $J \times J$ matrix $N$. Elements of this matrix, the scalars $n_{jj'}$, are constant across stores and represent the effect of a price change of item $j'$ on the sales of item $j$. The $J \times J$ diagonal matrix $A_k$ allows for store-specific own-elasticity components. The elements of this matrix are equal to $\lambda_k I$, with $\lambda_k$ the store-specific effect. Importantly, this effect is the same for all items within a category in store $k$. Hence, the own-elasticities ($n_{jj} + \lambda_k$) are allowed to vary across stores, while the cross-elasticities ($n_{jj'}; j \neq j'$) are similar across stores.

In this model the number of parameters is considerably lower than in a model with item and store-specific parameters. Furthermore, Hoch et al. (1995) include two additional restrictions to reduce the number of parameters. First, the own deal and own feature parameters are restricted to be constant across stores. Second, they restrict item-store intercepts to equal the sum of an item and a store intercept. They test for these restrictions and conclude that for their application, out-of-sample predictive validity is higher if these restrictions are imposed.

Hoch et al. (1995) obtain store-specific category price elasticities ($\eta_{ck}$) from a weighted aggregation of item-level price elasticities:

$$\eta_{ck} = w_k' (N + A_k) u$$  \hspace{1cm} (4.2)

with

$$w_k = J \times 1 \text{ vector weights to compute the category level price elasticity.}$$

The elements of the vector ($w_{jk}$) are the volume shares of item $j$ within the category.

These category-specific elasticities are explained from consumer- and competitive characteristics in the second step. Hoch et al. (1995) postulate the following relationship between elasticities and the predictor variables:

$$\eta_{ck} = X_k' \beta_c + e_{ck} \hspace{1cm} e_{ck} \sim N(0, \sigma^2)$$  \hspace{1cm} (4.3)

with

$$X_k = \text{predictor variables for store } k, \text{ including consumer and competitor characteristics,}$$

$$\beta_c = \text{parameter vector for category } c.$$
Chapter 4

- The variables associated with competing supermarket size and distance have weak and mixed effects.

Montgomery (1997)

Montgomery (1997) extends the model of Hoch et al. (1995). His major contributions are that he (i) integrates the estimation and prediction in one single step, (ii) imposes fewer restrictions on the elasticity structure, and (iii) shows how differences between stores can be exploited to increase profits.

Montgomery (1997) uses a Bayesian framework to estimate store-specific elasticities. This framework implies that the parameters are modeled as draws from a latent distribution, known as hyper-distribution. This distribution depends on consumer- and store characteristics.

The specification of the store-level sales model is:

\[
\ln \text{SALES}_{kt} = \alpha_k^* + H_k^* \text{PRICE}_{kt} + \Psi_k^* \text{DEAL}_{kt} + \Xi_k^* \text{FEAT}_{kt} + \epsilon_{kt}^* \tag{4.4}
\]

with

- \(\alpha_k^* = J \times 1\) vector with store-specific item intercepts. Elements of the vector are the \(\alpha_{jk}^*\), the intercept for item \(j\) in store \(k\),
- \(H_k^* = J \times J\) matrix of store-specific own- and cross-price effects,
- \(\Psi_k^* = J \times J\) diagonal matrix with on the diagonal item and store-specific feature effects \(\psi_{kk}^*\),
- \(\Xi_k^* = J \times J\) diagonal matrix with on the diagonal item and store-specific own-deal effects \(\xi_{kk}^*\),
- \(\epsilon_{kt}^* = J \times 1\) vector of random disturbances in store \(k\) in week \(t\). Elements of this vector are \(\epsilon_{jk}^*\), the disturbance for item \(j\) in store \(k\) in week \(t\).

The parameters of this demand system are stacked into a single vector:

\[
\theta_k^* = [\alpha_k^* \ \text{vec}(H_k^*)' \ \psi_{k11}^* \ldots \psi_{kll}^* \ \xi_{k11}^* \ldots \xi_{kll}^*]'
\]

This parameter vector is treated as a draw from a hyper-distribution. Each element of this vector \((\theta_k^*, i = 1, \ldots, K(K + 3))\) is modeled as a linear function of competitor and consumer characteristics of the store \((X_k')\). These variables are identical (same data) to the variables used by Hoch et al. (1995). The function is:
Existing models and store profile data

Hoch and al. (1995) estimate this model for 18 categories based on observations in 83 stores. They use 11 store profile variables to predict the category elasticities in the second step. These variables are:

**Consumer characteristics**
- the percentage of the population over 60 years of age;
- the percentage of the population with a college education;
- the percentage of households with five or more members;
- the log of median income;
- the percentage of houses with a value over $150,000;
- the percentage of women who work;
- the percentage of black- and Hispanic households.

**Competitor characteristics**
- the average distance in miles to the nearest five supermarket competitors;
- the distance to the nearest warehouse operation;
- the sales volume (as proxy for store size) of each store relative to the supermarket competition;
- the sales volume of each store relative to the nearest warehouse.

The consumer characteristics are the household characteristics in the store’s trading area. This trading area is determined by expanding a polygon around each store location that contains sufficient households to support the store’s ACV (All Commodity Value).

A comparison of the results across categories leads to the following conclusions about the effect of the relationship between price elasticities and consumer- and competitor characteristics:

- Education and housing value have, in general, a negative effect on price elasticity.
- The variables concerning ethnicity, family size and percentage of working women have a positive influence on price sensitivity.
- The variables related to income and age have mixed effects. On some categories they have a positive effect and on other categories a negative effect.
- The size of the nearest warehouse increases price sensitivity. The greater the distance to the nearest warehouse, the lower the sensitivity.
Existing models and store profile data

\[ \theta_{ki}^* = X_k^i \beta_{i}^* + \nu_{ik}^* \]  \hspace{1cm} (4.6)

with

\[ \beta_{i}^* = P \times 1 \text{ vector of effects of the predictor variables on } \theta_{ki}^*, \]

\[ \nu_{ik}^* = \text{random disturbance. The } \nu_{ik}^* \text{ for all } i \text{ can be stacked in a vector } \nu_k^*, \]

\[ [\nu_{ik}^* \ldots \nu_{ik}^*]' \]. Elements of this vector are \( \varepsilon_{jkt}^* \) the disturbance for item \( j \) in store \( k \) in week \( t \).

Montgomery imposes a set of restrictions on the parameters \( \beta_{i}^* \). The restrictions are that the \( \beta_{i}^* \) are the same for five clusters of parameters in the vector \( \theta_{ki}^* \). These clusters are the constants (parameters \( \beta_{i0}^* \)), own-price (\( \beta_{io}^* \)), cross-price (\( \beta_{ic}^* \)), feature (\( \beta_{if}^* \)) and deal parameters (\( \beta_{id}^* \)). These restrictions are less stringent than the restrictions Hoch et al. (1995) impose.

Model estimation in a Bayesian framework requires that prior distributions be imposed on the parameters. These prior distributions represent a-priori knowledge of the parameter estimates and serve as starting point for the estimation. The estimation results are the posterior distributions of the parameters. The prior distribution on \( \theta_{ki}^* \) is of special importance. This distribution determines to which extent stores have unique price elasticities. One extreme is that all stores have the same elasticities. This leads to a pooled model across stores. In the other extreme elasticities are similar to least squares estimates per store. Montgomery (1997) shows that a choice between these extremes is reasonable.

The results show that demographic effects are moderately predictive of the feature, own-price, and constant terms (\( R^2 \) varying between 0.14 and 0.25) and weakly predictive of cross-price and deal effects (\( R^2 \) of 0.02 and 0.05). Few of the estimated coefficients are significant. Montgomery (1997) claims that this can be improved upon by imposing stronger prior distributions. We would like to point out here that Montgomery did not include store characteristics. Store characteristics are likely to improve the parameter prediction, especially for the constant terms.

In a subsequent step, Montgomery (1997) shows how to employ store-specific price elasticities to increase store profits. He considers several possibilities with different restrictions. First, he optimizes prices under the restriction that the

---

\( ^2 \) See Montgomery (1997) for details on the prior distributions.
average category price and total category revenue per store remain constant. He shows that in his example (orange juice category) this strategy results in an average profit increase of 3.5 percent over the optimal uniform pricing strategy. Second, he performs the same optimization with price and revenue restrictions at the store level. This results in an additional profit increase of 0.6 percent (4.1 percent in total). Third, he compares a local marketing strategy with the current chain’s zone pricing strategy. This strategy implies that the chain uses different prices in different clusters of similar stores. The similarity is based upon competitive characteristics. He shows that assigning stores to these clusters based on the model results in a 3 percent increase over the current zone pricing strategy. Finally, he optimizes prices of individual stores in each zone (cluster) under the restriction that average prices and total revenue per cluster remain the same. This results in an 9.6 percent increase in profits over the optimal uniform pricing strategy.

Mulhern et al. (1998)

Mulhern et al. (1998) apply a two-step procedure to explain store-specific brand level own-price elasticities from brand characteristics and store profiles. They apply their model to a liquor store that is a monopolist in its market. The modeling approach is a two-step approach comparable to Hoch et al. (1995). In the first step they estimate store-specific brand level elasticities. In the second step they explain the estimated elasticities.

Mulhern et al. (1998) specify the following brand level model at the store level.

$$\ln \text{SALES}_{jkt} = \alpha_{kj}^{**} + \sum_{j'} \gamma_{kj'}^{**} \text{PRICE}_{jkt} + \delta_{jkt}^{**} \text{HDAY}_t + \varepsilon_{jkt}^{**} \quad (4.7)$$

with

- $HDAY_t = a$ dummy variable that indicates whether period $t$ is in the peak season (December),
- $\alpha_{kj}^{**} = a$ brand intercept for brand $j$ in store $k$,
- $\gamma_{kj'}^{**} = a$ parameter for the price effect of the price of item $j'$ on item $j$,
- $\delta_{kj}^{**} = a$ parameter for the effect in the peak season,
- $\varepsilon_{kj}^{**} = a$ disturbances for brand $j$ in store $k$ in period $t$. Mulhern et al. (1998) impose an AR(1) structure on these disturbances.
Existing models and store profile data

Model (4.7) is a sales response model with an exponential functional form. $\gamma_{jk}$ are the own price parameters, $\gamma_{jk'}$ ($j' \neq j$) are the cross price effects. Importantly, price parameters are not equal to the price elasticities in this model. Mulhern et al. (1998) evaluate own brand store-specific price elasticities at the average price. They calculate these elasticities ($\eta_{kij}^{**}$) by multiplying the price coefficient with the average price across the whole period for brand $j$ in store $k$ ($\bar{P}_{kj}$):

$$\eta_{kij}^{**} = \frac{\bar{P}_{kj}}{\gamma_{kij}}$$  (4.8)

Mulhern et al. (1998), in contrast with Hoch et al. (1995), do not impose a structure on the price parameters. The small number of competing brands considered allows this approach. Note that this is in fact a restriction on the number of brands.

The second step involves the prediction of these elasticities from the brand characteristics and store profiles. They estimate the following linear regression model:

$$\eta_{kij}^{**} = X'_k \beta^{**} + Z'_j \zeta^{**} + \nu_{kij}^{**}$$  (4.9)

with:

$Z_j^{**} =$ characteristics for brand $j$,
$\beta^{**} =$ vector of store profile effects,
$\zeta^{**} =$ vector of brand characteristic effects,
$\nu_{kij}^{**} =$ random disturbance for brand $j$ in store $k$.

Mulhern et al. (1998) model 14 liquor brands in 4 sub-categories sold in 35 stores in the USA. They use 3 store profile variables and 3 brand characteristics. Importantly, they do not include competitor characteristics since the liquor store is a monopolist in its market area. The store profile variables, all customer characteristics are:

---

3 Formally, the elasticities are equal to the price (not necessarily the average price) times the coefficient.
Chapter 4

- percentage of population within a one-mile radius of store \( k \) that is African-American;
- percentage of population within a one mile radius of store \( k \) that is Hispanic;
- median household income within a one mile radius of store \( k \).

The brand characteristics are:
- the market share for brand \( j \) for the entire period studied;
- percentage of time periods there was a price promotion for brand \( j \);  
- a dummy variable that indicates whether a brand is a premium brand.

Their analysis shows that the own brand price elasticity is higher in magnitude in (i) areas with lower concentrations of African-American consumers, (ii) market areas with higher incomes, (iii) more frequently promoted brands, and (iv) brands with a higher market share.

Campo et al. (2000) study the impact of consumer-, store- and competitor characteristics on the optimal space allocation of product categories within a store. They consider the effect of these factors on (i) the category’s relative attractiveness and (ii) total store sales. Campo et al. (2000) focus on the potential sales and do not include store-specific instrument effects.

The model of Campo et al. (2000) starts by expressing category sales \( (CSALES) \) as the product of store sales \( (SSALES) \) and category sales share \( (CSHARE) \). This results in the following identity for store \( k \) in period \( t \):

\[
CSALES_{ckt} = SSALES_{kt} \times CSHARE_{ckt} \tag{4.10}
\]

Next, they specify a model for each variable on the right side. The category share is modeled with an attraction model \( (MCI) \). The store sales are explained by a multiplicative model. The model for the category sales share is:

\[
CSHARE_{ckt} = \frac{A_{ckt}}{\sum_{c' \in C} A_{c'kt}} = \frac{\alpha_{0c}^{***} \prod_{s=1}^{S} X_{skt}^{\alpha_{sic}} \prod_{m=1}^{C} ASHARE_{mkt}^{\alpha_{mic}}}{\sum_{c' \in C} \alpha_{0c'}^{***} \prod_{s=1}^{S} X_{skt}^{\alpha_{sic}} \prod_{m=1}^{C} ASHARE_{mkt}^{\alpha_{mic}}} \tag{4.11}
\]
Existing models and store profile data

with

\[ A_{c_{kt}} = \text{the attraction of category } c \text{ in store } k \text{ in period } t, \]
\[ X_{s_{kt}} = \text{store profile variable } s \text{ for store } k \text{ in period } t, \]
\[ ASHARE_{m_{kt}} = \text{category space, the fraction of the store area allocated to category } m \text{ in store } k \text{ in period } t, \]
\[ \alpha_{0c}, \alpha_{sc}, \alpha_{2mc} = \text{parameters.} \]

Model (4.11) states that the category sales share depends on the category’s attraction relative to the total attraction of all categories. The attraction depends on the store profile variables (the consumer, competitor and store characteristics) and the area allocated to each category.

The model for total store sales is:

\[ SSALES_{kt} = \beta_{0}^* \prod_{s=1}^{S} X_{s_{kt}} A_{s_{kt}}^{\beta_{2}} \]

with

\[ SSALES_{kt} = \text{total sales for store } k \text{ in period } t, \]
\[ TA_{kt} = \sum_{c=1}^{C} A_{c_{kt}} \text{ the total attraction of all categories in store } k \text{ in period } t, \]
\[ \beta_{0}^*, \beta_{1s}, \beta_{2} = \text{parameters.} \]

Model (4.12) predicts total sales from the store profile variables and the total attraction of all categories together. The parameter \( \beta_{2}^{**} \) is of special importance as it represents the effect of the total attractiveness of all categories on total sales. Values of \( \beta_{2}^{**} \) higher than 1 imply that an attraction increase for a category leads to higher store sales. Each category benefits from these additional sales proportional to the category sales. If \( \beta_{2}^{**} = 1 \), then multiplying models (4.11) and (4.12) gives a multiplicative category level sales model that allows for cross-category effects. Values of \( \beta_{2}^{**} \) below 1 suggest there are diminishing returns on total attraction. In the extreme case \( \beta_{2}^{**} = 0 \), there is only between category substitution.

Campo et al. (2000) is the only study that includes store characteristics. They use the following variables:

**Store characteristics**

- store size.

53
Chapter 4

Consumer characteristics
- number of children per household (4 variables representing different classes);
- number of persons per household (4 classes);
- income (5 classes);
- marital status (4 classes);
- employment (5 classes);
- age (4 classes);
- ethnic background (2 classes);
- urbanization degree;
- number of people working but not living in the area.

Competitor characteristics
- own sales surface relative to competing supermarkets sales surface in the trade area;
- number of specialty stores in the trade area that compete with a category.

Six of the customer characteristics are captured by multiple variables representing different classes. Each of these variables indicates the population percentage in a class, for example the percentage of households without children, or with one child etc. One of the problems with these variables is that they are strongly correlated (multicollinearity). Campo et al. (2000) solve this by extracting four underlying uncorrelated factors: young families, less-well-to do, middle class, and single child. They use the scores on these variables in their analysis instead of the original variables.

Campo et al. (2000) model space elasticities as a function of store profiles. Space elasticities are proportional to category sales shares, which depend on store profiles. In fact this relation is a restriction. Store profiles are not allowed to affect parameter estimates in another way.

The model is applied to data from 55 stores of a European retail chain for two periods. For the category share model, they find that store profile variables have relatively strong effects on the attractiveness of the clothing-, luxury items-, dairy- (counter), meat- (counter), and fish categories. The influence of location characteristics on the attractiveness of groceries, meat (self-service), dairy (self-service), health and beauty is weak. For the total sales model, sales are higher in urban areas, for larger stores, and for stores with more competition. Campo et al. (2000) explain the counter-intuitive effect of competition from the fact that areas with more potential attract more competitors. The only factor (from the factor
Existing models and store profile data

analysis) that affects store sales is the middle class factor. Store sales are lower in areas with more middle class households.

Campo et al. (2000) compare the added value of optimizing space using store profiles with an optimization that does not use store profiles. Their results show that using profiles leads to 1 percent additional gross profits and 13 percent additional net profits.

Comparison and discussion

We described studies that assess the relationship between store profiles and marketing instruments. We compare these studies on the following aspects:

i. the modeling approach,
ii. which instruments are modeled at the store level,
iii. whether sales potential is included,
iv. the store profile variables used, and
v. modeling restrictions.

We summarize the comparison in Table 4.1.

With respect to the modeling approach we see that Hoch et al. (1995), Montgomery (1997), and Mulhern et al. (1998) all use store profiles to specify a sales response model that allows for store-specific parameter estimates. Montgomery (1997) estimates and explains these estimates in a single step. Hoch et al. (1995) and Mulhern et al. (1998) apply a two-step approach. They first determine parameter estimates and then explain these estimates from store profiles. The disadvantage of this approach is that uncertainty in the first step is not accounted for in the second step. The modeling approach of Campo et al. (2000) is different as they combine an attraction model and a sales model.

The marketing instrument in Campo et al. (2000) is category space. The other studies focus on price elasticities. Mulhern et al. (1998) consider promotional price elasticities. Hoch et al. (1995) and Montgomery (1997) study price without distinguishing promotional and regular prices. They include a dummy variable to account for temporary price discounts.

Potential sales are only considered by Montgomery (1997) and Campo et al. (2000). The model by Campo et al. (2000), unlike that of Montgomery et al. (1997), includes store characteristics. These characteristics are critical to determine potential sales.
Table 4.1 Comparison between the models in the marketing literature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modeling approach</strong></td>
<td>two-step</td>
<td>integrated</td>
<td>two-step</td>
<td>two models</td>
</tr>
<tr>
<td><strong>Store-specific instruments</strong></td>
<td>category level price elasticity</td>
<td>own price, cross price, feature, deal</td>
<td>own promotional price elasticity</td>
<td>area allocation across categories</td>
</tr>
<tr>
<td><strong>Potential sales</strong></td>
<td>no</td>
<td>yes</td>
<td>No</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Store profiles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Customer characteristics</strong></td>
<td>yes</td>
<td>yes</td>
<td>Yes</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Link between customers and stores</strong></td>
<td>distance-based trade area. distance depends on ACV.</td>
<td>distance-based trade area. distance depends on ACV.</td>
<td>households within 1 mile distance</td>
<td>survey-based trade area</td>
</tr>
<tr>
<td><strong>Competitor characteristics</strong></td>
<td>yes</td>
<td>yes</td>
<td>implicit (monopolist)</td>
<td>yes</td>
</tr>
<tr>
<td><strong>Restrictions</strong></td>
<td>Own price effect is the sum of an item-specific price effect that is constant across stores and a store-specific price effect that is constant across items.</td>
<td>Prior distribution on instrument effects. Store profile effects are the same for the own effect, cross price effects, feature, deal, and potential sales.</td>
<td>Number of competing brands is limited.</td>
<td>Store profiles influence area elasticity through their effect on sales share.</td>
</tr>
</tbody>
</table>

The disadvantage of models that estimate potential and instrument effects simultaneously is that the instrument effects may be biased. Estimating the sales potential requires the use of between store variation. This variation may be affected by unobserved retailer behavior. That is, retailer may use his marketing instruments based upon their expectations about sales. Stated differently, sales determine the use of the marketing instruments. If the store profile variables do not account for this behavior (i.e. we do not observe the behavior), biased parameter estimates result. In Chapter 6, we discuss this issue in detail. We show that biased estimates indeed result and we suggest an alternative approach.

Considering the store profile variables, we observe that all studies include consumer characteristics. Hoch et al. (1995) and Montgomery et al. (1997) use the same consumer characteristics, which correspond to the household characteristics in the store’s trade area. This trade area is defined as a polygon around the store that is large enough to support the total store sales (ACV). Mulhern et al. (1998) use household characteristics that refer to households that live within a one mile
radius. They use fewer characteristics compared to the other studies. Campo et al. (2000) use household characteristics in a trade area that is based on interviewing. The advantage of surveying is that the trade area definition is more exact than in the other studies.

The studies assign all consumers within the trade area or radius to the store. The problem with this approach is that consumers outside the target segment are considered as well. This leads to invalid characteristics. We expand this in the next section.

All studies account for competition. Mulhern et al. (1998) do this implicitly by considering a monopolist, Hoch et al. (1995) Montgomery et al. (1997), and Campo et al. (2000) include competitor characteristics in the models and thus explicitly account for competitor characteristics. Stores are identified as competitors if they are located within a certain distance or within the trade area. The characteristics are based on presence, distance, or relative sales volume/size. This definition may be wrong. We elaborate this in the next section.

All studies limit the number of parameters by imposing restrictions. These restrictions are needed because the number of observations per store over time is low (2-160 observations per store in the studies considered). This implies that it is not possible to estimate store-specific parameters for a large number of covariates. In Chapter 5, we show how daily data can be used to increase the number of observations. Hoch et al. (1995) restrict the own price effect to the sum of a brand specific effect constant across stores and a store-specific effect constant across brands. In addition, other parameters are assumed to be constant across stores. Montgomery (1997) imposes prior distributions on the parameters and requires the store profile’s effect to be equal for clusters of parameters. Mulhern et al. (1998) limit the number of brands. Campo et al. (2000) impose constant model parameters across stores. The store profiles determine the area elasticity through their effect on category sales share.

4.4 Data
Data on store profiles are needed to calibrate models described in the previous sections. In this section we describe the data on store-, competitor- and consumer characteristics available in the Netherlands. As the collection of consumer characteristics is more complex than collecting store and competitor characteristics, we spend more attention on this type of data.
4.4.1 Store- and competitor characteristics
Store- and competitor characteristics are related. Competitor characteristics are often defined as the store characteristics of competing stores. Frequently used variables for both types of characteristics include annual turnover, sales area, number of checkouts, store format and location (or distance to the store considered). Firms like ACNielsen (VNU), Elsevier and Locatus supply these data.

The definition of competition is critical for competitor characteristics. That is, if we want to obtain competitor characteristics we first need to know which stores compete. One possibility is to use distance or trade-area-based criteria like the ones used in the studies discussed in Section 4.3. The problem with this approach is that there is no guarantee that we will find the actual competitors. Some stores that are close to each other may not be competitors and vice versa. A better alternative is to use store choice data or to ask store managers to identify competition. The use of store choice data is more promising as these data are easier to obtain and more objective.

A more fundamental issue is that competition should be measured from the consumers rather than from the customer viewpoint. For an arbitrary customer one competing store may be dramatically closer (and more competitive) than for another customer. Spatial interaction models like Huff’s model account for these factors by modeling store choice through an attraction model that includes distance (see e.g. Ghosh et al., 1994, pp. 305-313). The very detailed data needed to calibrate these models at lower levels of aggregation (e.g. brand level) are not available in the Netherlands.

4.4.2 Consumer characteristics
When we consider consumer characteristics, we must first decide which consumers to focus on. We divide consumers into potential- and actual customers who may have different characteristics. For local marketing we consider two

---

4 Alternatively one might want to account for the competitor’s marketing instruments. We do not consider this possibility as these data are not (yet) available.
5 One might argue that this also yields competitor characteristics (e.g. percentage that visits a certain store). We consider this as an intermediate variable between competitor and customer characteristics.
6 This can be refined by considering the degree to which customers are potential customers (likelihood of becoming an actual customer) or actual customers (is the store primary or secondary store, etc.).
possible target groups. The first option is to focus on actual customers. In this case it is necessary to identify the store’s customers. The second option is to focus on both potential- and actual customers. In this case we may use consumer characteristics in the trade area, just as in the studies in the previous section. However, there are some serious concerns with this approach. First, we assume that everyone in the trade area is an actual- or potential customer, which is unlikely. Second, a correct definition of the trade area is critical. A wrong definition results in incorrect consumer characteristics. Third, consumers are assigned to a store based on their residential location. This excludes the possibility of transient consumers, i.e. consumers who do not live in the area. Campo et al. (2000) attempt to solve this by including the number of people that work but do not live in the area\(^7\).

The choice whether to consider potential customers or not determines which data collection method is suitable. A valid method measures the characteristics of the selected target segment. We consider the methods’ data collection location and possibility to identify customers for comparison. Other characteristics we consider are the availability of the data and the number of characteristics that can be included. First, we describe the methods using these characteristics: see Table 4.2. Next, we compare individual- and aggregate household data in the Netherlands.

We distinguish two possible data collection locations, i.e. (1) at the checkout, and (2) at the consumer (household) address. At the checkout, data can be collected using customer cards or by asking for the customer’s zip code. Customer cards are used by the store to reward their customers. The customer receives a card upon registration and is supposed to present their card at the checkout. This gives the store insight into which products the customer buys. Additional customer characteristics can be obtained from a survey (e.g. at registration) or other data sources.

Asking for zip-codes is a less sophisticated way to collect data at the checkout. The zip-code reveals the customer’s residential area. The store may add additional characteristics from an external data source (e.g. survey, zip code level database).

\(^7\) Comparable data is available in the Netherlands from registrations on the number of people working in an area (suppliers in the Netherlands are KvK, LISA)
Chapter 4

Table 4.2 Characteristics of methods to collect consumer data

<table>
<thead>
<tr>
<th>Method</th>
<th>Data collection location</th>
<th>Possibility to identify customers</th>
<th>Availability</th>
<th>Number of characteristics that can be included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer cards</td>
<td>checkout</td>
<td>yes</td>
<td>limited</td>
<td>few</td>
</tr>
<tr>
<td>Checking zip-codes</td>
<td>checkout</td>
<td>yes</td>
<td>limited</td>
<td>few</td>
</tr>
<tr>
<td>Survey data</td>
<td>household address</td>
<td>yes</td>
<td>limited</td>
<td>many</td>
</tr>
<tr>
<td>Individual household data</td>
<td>household address</td>
<td>yes</td>
<td>widely</td>
<td>many</td>
</tr>
<tr>
<td>Aggregated household data</td>
<td>household address</td>
<td>no</td>
<td>widely</td>
<td>many</td>
</tr>
</tbody>
</table>

At the address level, data can be collected through a survey or can be purchased from a third party. Third parties include firms whose business it is to collect and sell data. These data are available either at the individual household-level or are aggregated across households. We compare individual- and aggregate household data supplied by third parties in the Netherlands below.

Customer identification is critical if we want to focus on actual customers. Without correct identification wrong characteristics may result. Data collection at the checkout guarantees that all respondents are customers. Still, a bias may result if customers do not use their card or refuse to reveal their zip-code. At the address level customers may be identified by asking them which stores they visit. Thus it is possible to identify stores with individual household data. On the other hand, aggregate household data only reveals the percentage of respondents who visit which store. Therefore, we cannot determine the characteristics of these households.

There is no need to identify individual customers if we want to know both current- and potential customer characteristics. In this case it suffices to determine the store’s trade area and corresponding consumer characteristics. Store choice data may be used to determine the trade area (if available), or alternatively a distance-based criterion. Examples of distance-based criteria include (i) considering a radius around the store, and (ii) expanding polygons around stores such that they are sufficiently large to support the total store sales.

Availability refers to the extent to which it is possible to obtain data. Data obtained from customer cards, asking for zip-codes and surveys are not widely available. These data are collected by a limited number of stores and the owner
Existing models and store profile data

(often the retailer) is not likely to share the data. By contrast, data offered by a third party are widely available. Third parties collect data on a large scale and are willing to sell the data to everyone.

The amount and type of consumer characteristics collected varies dramatically across the methods. In general, data collection at a household’s address offers more opportunities to include characteristics than data collection at the checkout. The situation at the checkout (or at the customer card registration) does not allow many questions. On the other hand, the advantage of collecting data at the checkout is that actual purchase behavior is observed, though only for the store considered. At the address level, communication is the only option to collect purchase behavior. Stated purchase behavior may differ from actual behavior.

One may combine data from a survey or third party with data collected at the checkout. Both individual- or aggregate household data can be used, although both approaches have their disadvantages. Using individual household data implies that data about the same customer are added which will not be available for all customers. Using aggregate data implies that aggregate characteristics are assigned to individual customers. This assumes that individual customers have the aggregate characteristics (or at least on average). It is obvious that this assumption is easily violated.

**Individual household data**

Individual household data are offered by different firms in the Netherlands. Cendris, Claritas, and WegenerDM are important suppliers. These suppliers offer large databases that cover a broad range of individual household characteristics. We study the differences between these databases and collected them into Table 4.3.

All suppliers collect their data mainly through written customer surveys. These surveys are quite extensive and cover a variety of topics including demographics, income, insurances, durables, leisure time, cars, housing, media use, and store choice. These databases are built by combining data from repeatedly sending the same survey to different parts of the population. The databases of Claritas (CCI database) and Cendris cover 25 percent of the households. The database of WegenerDM is considerably smaller and covers 16 percent of households.
Chapter 4

Table 4.3 Differences between suppliers of individual household data in the Netherlands

<table>
<thead>
<tr>
<th></th>
<th>Cendris</th>
<th>Claritas (CCI)</th>
<th>WegenerDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of households</td>
<td>25%</td>
<td>25%</td>
<td>16%</td>
</tr>
<tr>
<td>Store choice</td>
<td>food</td>
<td>food</td>
<td>food</td>
</tr>
<tr>
<td></td>
<td>drug</td>
<td>liquor</td>
<td>clothes</td>
</tr>
<tr>
<td></td>
<td>clothes</td>
<td>drug</td>
<td>department stores</td>
</tr>
<tr>
<td></td>
<td>department stores</td>
<td></td>
<td>home improvement</td>
</tr>
<tr>
<td>Sponsoring</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

All suppliers identify stores visited from a question about visited *chains*. Households are assigned to the nearest store of a chain. This is incorrect for households that do not visit the nearest store of a chain (e.g. the store close to the work address). It would be better to identify the visited *stores* directly, though this is more difficult to implement.

The three suppliers include different store types. Claritas is the only supplier who includes liquor stores and WegenerDM the only supplier who includes home improvement. Importantly, Claritas and WegenerDM distinguish between primary and secondary food stores. Cendris asks only for the most visited store. Additionally, for a number of categories Cendris and Claritas ask whether respondents visit specialty stores.

Claritas and WegenerDM allow sponsors to include questions. In this way the sponsor may obtain exclusive information on behavior-related characteristics such as product use, purchase quantities, and purchase frequency. The alternative, to predict these variables from other characteristics often results in weak predictive validity.

*Aggregated household data*

Aggregated household data summarize characteristics in an area. We classify the aggregation possibilities to the area definition used. The three definitions are (i) zip codes, (ii) neighborhoods and districts, and (iii) shopping districts. We explain these methods and suppliers below: see Table 4.4.
Existing models and store profile data

Table 4.4 Differences between suppliers of aggregated household data in the Netherlands

<table>
<thead>
<tr>
<th>Area definition</th>
<th>Claritas</th>
<th>WegenerDM</th>
<th>Cendris</th>
<th>Experian</th>
<th>Cherridata</th>
<th>Statistics Netherlands</th>
<th>Locatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest aggregation level</td>
<td>PC6</td>
<td>PC6</td>
<td>PC6</td>
<td>PC6</td>
<td>PC7</td>
<td>PC4 / neighborhoods and districts</td>
<td>shopping district</td>
</tr>
<tr>
<td>Store choice</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

Zip-code level data

The zip code is the most frequent aggregation level used. Zip codes were introduced by the postal services to facilitate mail delivery. The Dutch zip codes consist of a combination of four digits and two characters (e.g. 1234 AB). On average a full zip code covers 17 households. The aggregation level defined by the full zip code is called the PC6 level. PC refers to PostCode (zip code in Dutch) and 6 to the number of positions used. Using fewer positions identifies households in a larger area. Thus, the PC5 level refers to addresses with identical first 5 positions and the PC4 level refers to addresses with the same four digits. On average there are 200 households at the PC5-level and 1500 at the PC4-level. Adding a character (or number) to the original zip code makes it possible to identify households at a more disaggregate level than the PC6 level. This aggregation level is called the PC7 level.

Claritas, Cendris, WegenerDM, and Cherridata all offer PC4, PC5, and PC6 level databases. Cherridata is the only supplier that offers a database at the PC7 level. These databases are based on individual household data. In addition, all suppliers offer a geodemographic segmentation system. These systems assign each zip code to a lifestyle segment.

Experian is another supplier of a geo-demographic segmentation system. This segmentation system, called Mosaic, is based on data from different sources and is available at all zip-code levels. Store choice is not included.

Statistics Netherlands (CBS in Dutch) offers data at the PC4 level. Municipal- and car administration databases are the main sources. The data include the number of inhabitants, age, nationality, urbanization degree, address density, income and car possession. Importantly, store choice is not available. The advantage of the data supplied by Statistics Netherlands is that the price is low.
The disadvantages are that the data are not detailed (high aggregation level) and that a limited number of characteristics is available.

**Neighborhoods and districts**
The categorization into *neighborhoods* and *districts* is very common within municipalities in the Netherlands. These areas are defined such that a district contains multiple neighborhoods. On average there are 575 households in a neighborhood and 2400 households in a district. One problem with this area definition is that the number of household varies dramatically across neighborhoods and districts. Within neighborhoods, the number of households can vary from 0 to 10,000. Statistics Netherlands offers data at these aggregation levels. The variables in Statistics Netherlands neighborhood and district data and the PC4 level data are identical.

**Shopping centers**
Locatus offers customer information at the shopping center level. This approach is unique in that it defines a common customer profile for a whole shopping center. This is unattractive if a store-specific profile is needed.

### 4.5 Summary and conclusions
In this chapter we studied models and data for local marketing. We noted that the models in the marketing literature relate store profile data to the effects of marketing instruments, but these models are not fully implemented in practice.

We illustrated models in practice by explaining the services of ACNielsen in the Netherlands that relate store profiles to store sales. These services are *FamilyTrack*, *Local Marketing*, and *Assortman*.

We discussed four models in the marketing literature noticing the following weaknesses:

- Parameter restrictions have to be imposed as the number of observations is usually low.
- Estimating sales potential requires the use of between store variation which may lead to biased estimates.
- Competitor characteristics are based on distance and not on choice criteria. Hence, competition is not considered from the consumer point of view.
Existing models and store profile data

- Consumer characteristics are based on consumers living within a radius or the trade area. None of the studies distinguishes between actual- and potential customers.

We provide solutions to the first two weaknesses in Chapters 5 and 6. In the second half of this chapter we discussed the possibilities of obtaining data for local marketing and the data suppliers in the Netherlands. For the competitor data, we noted that it is better to use store choice data to identify competition. Alternative models that measure competition from the customer point of view require data not available in the Netherlands. For consumer data there are two options. We can focus either on actual customers or on actual and potential customers. Finding valid consumer characteristics for actual customers is easier than for actual and potential customers.