Chapter 5
Sales Decomposition Within and Across Categories Using Daily Data for a Single Store

5.1 Introduction
In this chapter we develop and apply a model to estimate store-specific marketing instrument effects using daily data for a single store. In this way we implicitly account for differences between stores.

Most store-level sales models use weekly data and multiple stores. If we use these data to estimate models with unique estimates per store we often have insufficient observations to estimate the effects of the variables that affect sales. The use of daily data increases the number of observations and hence we can include more instruments.

We decompose an item’s promotion effects into within- and cross category effects. For many managers it is critical to consider interdependencies between products offered in the market place. For example, a multi-product retailer should know how items in different categories interact. Many manufacturers and retailers have adopted category management to accommodate dependencies between multiple brands (e.g. cannibalization). They should also understand dependencies between categories because this may influence the nature of the marketing mix. However, little is known about category dependencies (see e.g. Manchanda, Ansari, and Gupta, 1999).

Neslin (2002, pp. 62-63) points out that none of the existing sales promotion decomposition models answer the question how an item’s sales promotion affects the sales of all following four components:

i. a change in the own-item sales of the item promoted;
ii. a change in sales of other items belonging to the same category (cross-item effects);
iii. a change in sales in other periods (cross-period effects);
iv. a change in sales of items in other categories (cross-category effects). These effects can be either positive (complementary) or negative (substitution).

Most existing decomposition studies consider elasticities. Sales elasticities are decomposed in terms of category incidence, brand choice and quantity elasticities. However, the elasticity percentages are not equal to unit sales percentages, and only the latter have managerially relevant implications (Van Heerde, et al. 2003b). Van Heerde, et al. (2003a) consider three of the four sales components mentioned by Neslin (2002, pp.62-63). They show how the own-brand unit sales effect of a promotion can be decomposed into cross-brand-, cross-period-, and category-expansion effects.

We expand their framework to include effects across categories; component (iv). We consider how the evaluation of a promotion (prices) changes with the inclusion of cross-category effects. We separate these effects into effects on other items of the promoted brand and on items of other brands. The first effect affects both the manufacturer’s and the retailer’s total sales. The second effect only affects the retailer’s sales, assuming no competitive reactions. In this chapter we present an approach to obtain store-specific marketing instrument effects from daily data for one store, we provide an approach to decompose the sales change into effects on items of the same brand and other brands across categories.

This chapter is organized as follows. In Section 5.2 we provide an overview of related research on cross-category effects, decomposition models and models for daily data. In Section 5.3 we propose our modeling approach. The data sets used to estimate the models are described in Section 5.4. In Section 5.5, we provide empirical results. We present our conclusions and discuss managerial implications in Section 5.6.

5.2 Literature review
First, we provide an overview of the studies on cross-category effects; second, we discuss the (single) study that uses daily data to model sales promotions; and finally, we discuss the existing literature on the decomposition of sales promotions.

5.2.1 Cross-category effects
Several studies consider cross-category relationships, some based on household data while other studies use store data.
Sales decomposition using daily data for a single store

*Household-level studies*
ErDEM (1998) applies the theory of signaling to study the relation between umbrella brands in two product categories. She finds strong support for positive cross-category effects. Erdem and Sun (2002) investigate the effects of advertising and promotions on items of umbrella brands. They show that information-based sales promotions (especially couponing) have a relatively large impact on umbrella brand items in other categories. Hruschka, Lukanowicz and Buchta (1999) estimate a multivariate binomial logit model to estimate cross-category dependencies. For 73 categories they find significant interaction effects for 5 percent of all possible category pairs. With a p-value of 0.05, this relative frequency is consistent with what might be expected to occur by chance. Russell and Petersen (2000) specify a conditional choice model for a basket of goods. They find modest cross-category effects for four different paper goods categories. Manchanda, Ansari and Gupta (1999) find cross category effects for pairs of complementary product categories (cake mix, cake frosting, fabric detergent, and fabric softener).

*Store-level studies*
Doyle and Saunders (1990) consider cross-category effects of advertising on sales for a variety store chain. They project that an advertising strategy that incorporates cross-category effects will result in a 40 percent profit increase. Mulhern and Leone (1991) use weekly store-level scanner data to estimate promotion effects between complementary categories (cake mix and frosting). Their findings also suggest that promotions in one category enhance sales in the other.

Walters (1991) investigates cross-category, intra- and inter-store substitution effects of brands within two predefined sets of related categories (cake mix and cake frosting, and spaghetti and spaghetti sauce) and between two stores. His findings reveal significant inter-store complementary and substitution effects but show little evidence of intra-store effects.

The conclusion based on previous studies is that cross category effects matter but are not always easy to isolate. Cross category effects are found especially in predefined sets of categories for which interactions can be expected.
5.2.2 Modeling daily data

By using daily data instead of weekly data, we have a larger number of observations and we can specify and estimate models for a single store that include many instruments. An additional advantage of daily data is that we avoid possible ambiguities with measures of price and feature. That is, we avoid possible measurement errors in weekly data caused by midweek changes in price and feature. Marketing instruments may not be aligned with the observation frequency.

In the marketing literature, we know of only one study that uses daily scanner data. Kondo and Kitagawa (2000) model category sales for a single store. They use the Kalman filter to separate category sales into a trend, a day-of-the-week effects, and an irregular component. The irregular component is modeled as a function of price promotions. They do not allow the promotion effects to differ by day nor do they accommodate different disturbance variances (heteroskedasticity) between days. By contrast, we expect promotion effects and disturbance variances to be proportional to average sales for each day of the week. Importantly, Kondo and Kitagawa (2000) do not decompose sales promotion effects and they do not consider cross-category effects.

5.2.3 Elasticity decompositions

The elasticity decompositions based on household level data include Gupta (1988), Chiang (1991), Chintagunta (1993), Bucklin et al. (1998), and Bell et al. (1999). They use the idea that sales elasticity is the sum of purchase incidence-, purchase quantity-, and brand choice elasticities. Bell et al. (1999) find that on average the sales promotion elasticity is mostly due to brand choice elasticity (74 percent), leaving only 26 percent for the purchase incidence and quantity elasticities. Van Heerde et al. (2003b) show that these results cannot be interpreted substantively. In fact they find that on average only 33 percent of the unit sales increase due to a promotion is attributable to brand switching.

Pauwels et al. (2002) study the long-term effects of a price promotion. They find that negative post promotion effects tend to cancel the brand choice elasticities (secondary demand), but do not cancel the category incidence- and quantity elasticities (primary demand).
Sales decomposition using daily data for a single store

5.2.4 Unit sales decomposition
Van Heerde et al. (2003a) decompose the unit sales promotion effect based on store-level scanner data into three sources: cross-brand effects, cross-period effects, and category expansion effects. They find that on average about 1/3 of the own-brand sales effect is due to the cross-brand effect, about 1/3 is due to the cross-period effect, leaving about 1/3 for the category expansion effect. Thus, they find a primary demand effect of 2/3 that corresponds well with the unit sales result inferred from the household model elasticities (Van Heerde et al., 2003b).

Van Heerde et al. (2003a) provide two extensions (extended decompositions). One extension decomposes the within-store category expansion effect into a market expansion effect (the category expansion effect net of within-brand cross-store effects) and a cross-store effect (changes in the focal brand sales in competing stores). The other extension is a decomposition of the cross-item effect into within-brand (cannibalization) and between-brand effects.

5.2.5 The decomposition model by Van Heerde et al.
Van Heerde et al. (2003a) use store-level scanner data to decompose the effect of sales promotions. For the standard decomposition the criterion variables are:

i. own-brand sales in week \( w \) (\( OBS_w \)) to estimate the own-brand effect,
ii. sales of all other brands in week \( w \) (\( CSB_w \)) for the cross-brand effect,
iii. category sales before and after week \( w \) (\( CSW_w \)) for the cross-period effect,
iv. category sales before, during, and after week \( w \) (\( CST_w \)) for the category expansion effect.

They estimate the models pooled across stores. Since stores differ in size, they transform the store sales to obtain equal promotion effects across stores proportional to average category volume. Thus, each sales variable is divided by the average category sales in the store (\( CS_k \)). If \( S_{kjw} \) is unit sales of brand \( j \) in store \( k \) in week \( w \), they define the criterion variables for the standard decomposition as:

\[
OBS_{kjw} = \frac{S_{kjw}}{CS_k}, \quad CBS_{kjw} = \sum_{j' \neq j} \frac{S_{kj'w}}{CS_k},
\]

\[
CSW_{kw} = \sum_{s=0}^{T} \sum_{j} \sum_{j' \neq j} \frac{S_{kj'w+s}}{CS_k}, \quad CST_{kw} = \sum_{s=2}^{T} \sum_{j} \sum_{j' \neq j} \frac{S_{kj'w+s}}{CS_k},
\]
where $J$ is the number of brands, $T$ is the number of lagged periods, and $T^*$ is the number of lead periods. By construction, the criterion variables are related to each other in the following way:

$$OBS_{kjw} = CST_{kw} - CBS_{kjw} - CSW_{kw} \quad (5.1)$$

Ignoring the common divisor ($CS_k$) this equation says that current own-brand unit sales is total category sales in a time window, minus current cross-brand sales, minus total category sales in the same time window excluding week $w$. The standard approach consists of estimating linear regression models for $OBS_{kjw}$, $CBS_{kjw}$, and $CSW_{kw}$ as a function of own-brand price index (actual price divided by regular price) and other variables. By construction, the effects for $CST_{kw}$ are implied by and can be derived from the other effects.

Van Heerde et al. (2003a) define four mutually exclusive types of price index variables depending on the support type. These support options are: (i) feature-only, (ii) display-only, (iii) feature and display, or (iv) no support. For each of the four promotion variables they obtain separate decompositions.

The regression models for each criterion variable include all relevant covariates to avoid possible biases due to missing variables. For example, the covariates include price (index) variables of other brands in the category for the current period, even though the cross-brand effects of interest are obtained from the effect of own-brand price indices on the $CBS$ variable. Similarly, lead and lag variables are included, even though the relevant cross-period effects come from the $CSW$ variable. Other covariates include regular prices, support variables without a price discount, and weekly dummies. The covariates for the other brands are defined as averages across the brands to minimize the number of covariates in the model. By including exactly the same covariates in each equation, the effects for $CST_{kw}$ are indeed implied by the other effects.

In general, the regression model for $OBS_{kjw}$ can be written as:

$$OBS_{kjw} = \sum_{l=1}^{4} \alpha_{ljOBS} PI_{lkw} + \sum_{p=1}^{P} \delta_p X_{pkjw} + u_{kjw} \quad (5.2)$$

with

$PI_{lkw} = $ price index variable (index) for brand $j$ in store $k$ in week $w$, where the index $l$ indicates the four different support conditions,
Sales decomposition using daily data for a single store

\[ X_{pjkw} = \text{observation for covariate } p, p = 1, \ldots, P \text{ on brand } j \text{ in store } k \text{ in week } w, \]

\[ u_{jkw} = \text{disturbance term for brand } j \text{ in store } k \text{ in week } w. \]

The parameter \( \alpha_{\text{OBS}} \) in (5.2) is the effect of the price index with support type \( l \) on own-brand unit sales divided by average category sales (\( CS_k \)). From the same regression models for \( CBS_{kjw} \) and \( CSW_{kw} \), the following parameters are estimated:

- \( \alpha_{\text{CBS}} \): the cross-brand effect (criterion variable \( CBS_{kjw} \)), and
- \( \alpha_{\text{CSW}} \): the cross-period effect (criterion variable \( CSW_{kw} \)).

And based on (5.1) it follows that

\[
\alpha_{\text{CST}} = \alpha_{\text{OBS}} + \alpha_{\text{CBS}} + \alpha_{\text{CSW}} \tag{5.3}
\]

where \( \alpha_{\text{CST}} \) is the category expansion effect (criterion variable \( CST_{kw} \)).

Thus, the own-brand effect equals the category expansion effect minus the cross-brand effect minus the cross-period effect. For estimation one of the equations for these four variables is deleted because it is redundant. Van Heerde et al. (2003a) use (5.3) to compute the relative contribution for each component as follows:

- fraction cross-brand effect: \( \frac{-\alpha_{\text{CBS}}}{\alpha_{\text{OBS}}} \)
- fraction cross-period effect: \( \frac{-\alpha_{\text{CSW}}}{\alpha_{\text{OBS}}} \)
- fraction category expansion effect: \( \frac{\alpha_{\text{CST}}}{\alpha_{\text{OBS}}} = 1 - \frac{\alpha_{\text{CBS}}}{\alpha_{\text{OBS}}} - \frac{\alpha_{\text{CSW}}}{\alpha_{\text{OBS}}} \)

5.3 Modeling approach

The modeling approach proposed by Van Heerde et al. (2003a) is our starting point. We explain how we adapt this model in Section 5.3.1. These modifications have consequences for the variable definitions, variable transformations, and model specification. In Section 5.3.2 we present the within-category and cross-
category decompositions and the models needed to determine these decompositions.

The modeling approach can be divided into three stages and nine steps. In the first stage, we transform the model variables. In the second stage, we use the transformed variables to estimate the instrument effects. In the third stage, we calculate the decomposition effects from the modeling results. We summarize the modeling steps in Table 5.1. In the following sections we refer to the steps in this table.

Table 5.1 Modeling steps

<table>
<thead>
<tr>
<th>Step</th>
<th>What</th>
<th>Why</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transformation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Multiplying unit sales by the average price across the sample</td>
<td>To obtain comparable units across categories</td>
</tr>
<tr>
<td>2</td>
<td>Dividing actual price by the average price</td>
<td>To obtain comparable price variables</td>
</tr>
<tr>
<td>3</td>
<td>Standardizing sales per day</td>
<td>To eliminate day-of-the-week effects</td>
</tr>
<tr>
<td>4</td>
<td>Filtering sales using the HP filter</td>
<td>To create fluctuations in sales not attributable to trend or seasonality</td>
</tr>
<tr>
<td>5</td>
<td>Filtering price index and feature multicollinearity problems</td>
<td>Same as in 4 for price and feature</td>
</tr>
<tr>
<td><strong>Estimation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Model estimation using the transformed variables</td>
<td></td>
</tr>
<tr>
<td><strong>Decomposition</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Multiplying regression coefficients by the average standard deviation per day</td>
<td>To obtain average daily effects in terms of sales</td>
</tr>
<tr>
<td>9</td>
<td>Using significant regression coefficients to find decomposition effects</td>
<td>To obtain reliable decomposition results</td>
</tr>
</tbody>
</table>

5.3.1 Modifying the model of Van Heerde et al. (2003a)

We modify the modeling approach of Van Heerde et al. (2003a) such that it is applicable to daily data for one store on multiple categories. The modifications

1 We conduct a sensitivity analysis to determine how the results change for different p-values.
are related to (i) variable definitions, (ii) trend and seasonality, and (iii) model specification issues.

(i) Variable definitions

Sales (step 1)

To accommodate cross-category effects, it is critical that we use criterion variables expressed in comparable units. For example, the unit for washing detergents is kilos but for fabric softener it is liters. We create comparable variables and define a revenue like variable in terms of unit sales at the item level times the average price for the item across the sample (called SAP). We use average price instead of actual prices to avoid artificial correlation problems between, for example the criterion variable and the price index. Hence the unit sales, for a given day are multiplied by a constant which only has a scaling effect.

Covariates

We define all covariates at the item level. We focus on price index effects and consider feature as a covariate.

Price

For many items it is difficult to distinguish between promotional- and regular price changes. Available algorithms to infer regular prices do not apply due to excessive price fluctuation. We proceed as follows. Price indices are calculated using the average price across the year as a reference price (step 2 in Table 5.1). In this way we measure all price effects in terms of a comparable price index.

Feature

Feature is the only support variable for which we have information. The pattern of variation in feature and price does not allow us to estimate separate effects for feature supported- and non-supported prices. Estimation of such effects requires that both supported and non-supported prices vary sufficiently, which is not the case. Therefore, price and feature are included as separate variables, and we do not allow for a separate interaction effect.

Correlation between predictors (step 6)

The store uses marketing instruments simultaneously for some items. For instance we see that price decreases for different items of the same brand are identical, or that for one item a decrease is always supported by feature. In the most extreme case instruments are perfectly correlated, the effects cannot be estimated
separately. In less extreme cases estimates may have high variances and incorrect signs.

We detect multicollinearity problems using a variance decomposition of the estimators based on the characteristic roots and vectors of the predictors’ moment matrix (see Judge et al., 1985, pp. 902-904, and Belsley et al. 1980). Belsley et al. (1980) show that this approach is superior to alternative approaches including inspection of the covariance matrix of predictors. Variance proportions of all estimates for roots with a condition index\(^2\) exceeding 15 are evaluated. A variance proportion for two or more predictors exceeding 0.50 indicates multicollinearity. Both threshold values are at the lower end (most prudent) of commonly suggested values\(^3\) (see Belsley et al., 1980, pp. 100-104, Greene, 2000, p. 258, and Hair et al., p.153).

Another issue is the case where instruments are often used simultaneously but not so frequently that severe multicollinearity results. In these cases the information present in the simultaneous use of instruments is not used fully for estimation. The separate instrument estimate for one of these items is based on the relation between this instrument and sales, both corrected for the other item’s instrument. However, from a managerial viewpoint the effect of joint instrument use is relevant. We detect these cases from the predictor variables’ covariance matrix, and we treat them as collinear variables.

We choose the following approaches to overcome collinearity problems: (i) we omit the feature variable if price and feature are (almost always) used together, and (ii) we create new price variables if prices are used simultaneously. The new price variables are constructed such that we have: (i) a common price variable that equals the common price if the prices are the same and the annual average price of both items if not, and (ii) separate price variables for each item that takes the value of the item’s price if prices are not the same, and the annual average price of both items if prices are the same. We explain this in detail in Appendix A.

(ii) Trend and seasonality
Store-level studies typically use weekly data from multiple stores. Time series components common to all stores are accommodated with weekly dummies to

---

\(^2\) The condition index is the ratio of the largest and smallest eigenvalues of the moment matrix.

\(^3\) We recognize that the degree to which multicollinearity is a problem depends on the number of observations (see e.g. Leeflang et al., 2000, p. 358).
account for seasonality and missing variables such as coupons and television advertising (e.g. Wittink et al. 1988 and Van Heerde et al. 2003a). With data for only one store the use of dummy variables for each of the time periods is not possible. Hence, we apply a different approach.

It is common to decompose a time series in period $t$ into a trend- ($T_t$), a seasonal- ($SEAS_t$) and an irregular component ($I_t$). For an additive model this means that sales on day $t$ ($S_t$) is expressed as:

$$S_t = T_t + SEAS_t + I_t$$ (5.4)

With data for only one store, we account in two subsequent stages for the day-of-the-week effects (step 3) and the trend in the data (steps 4 and 5) so that the irregular component becomes the basis for inferring marketing mix effects.

**Day-of-the-week effects (step 3)**

Systematic variation in sales between days of the week needs to be accommodated. For example, stores sell more on Saturdays than on Mondays. A standard approach is to model it in a deterministic way. For example, one could include dummy variables for days of the week. However, there are two additional differences between the days that require another modeling approach.

First, we expect marketing instruments to have effects proportional to sales on a typical day. For example, if a product on average sells six times as much on Saturday than on Monday, we expect a price decrease on Saturday also to have six times the impact of a price decrease on Monday.

Second, we expect that the disturbances have variances that depend on the day of the week (heteroscedasticity). We assume this variance to be proportional to the average daily sales. Such proportionality assumptions are implicit in multiplicative models (e.g. Wittink et al. 1988). However, we require an additive model to obtain a consistent additive sales decomposition (Van Heerde et al., 2003b). Our solution is to transform the sales variable such that effects of weekdays are modeled implicitly. That is, allowing for day-of-the-week effects is equivalent to subtracting average sales for a given day (e.g. average Saturday sales) from every observation on that day of the week. Then, proportionally equal effects of marketing instruments and disturbances are obtained by dividing the difference from the average by the standard deviation for each day.
Hence the sales of brand \( j \) from category \( c \) (\( j_c \)) on day \( d \) in week \( w \) are transformed into:

\[
ZSALES_{j,dw} = \frac{SAP_{j,dw} - \tilde{SAP}_{j,d}}{\hat{\sigma}_{SAP_{j,d}}} 
\]

with

- \( ZSALES_{j,dw} \) = standardized sales for item \( j \) in category \( c \) (\( j_c \)) on day \( d \), \( j_c = 1, \ldots, J_c \), \( c = 1, \ldots, C \), \( d = 1, \ldots, 6 \), \( w = 1, \ldots, W \),
- \( SAP_{j,dw} \) = sales for item \( j_c \) measured as unit sales × average price on day \( d \) of week \( w \),
- \( \hat{\sigma}_{SAP_{j,d}} \) = estimated standard deviation of the sales for item \( j_c \) on weekday \( d \),
- \( \tilde{SAP}_{j,dw} \) = \( SAP_{j,dw} \) average sales across all observations on day \( d \).

The index \( d \) indicates the day of the week and is defined for six days. The first five days refer to Monday through Friday. The sixth day is defined as the total sales on Saturday and Sunday. In most cases this is equal to the sales of Saturday as the store opens only on a few Sundays per year. In these cases the store is closed on the preceding Saturday.

**Detrending (steps 4 and 5)**

Detrending the data implies that we remove the long-term fluctuations from the data. For this purpose Abraham and Lodish (1987, 1993) use a Hanning smoother and several moving averages. We prefer and use the Hodrick Prescott filter, because it removes the long term trend without affecting the short term fluctuations too much. Lack of this property is the so-called distortion effect and it has been shown that this effect is minimal for this filter (Pedersen, 2001).

The Hodrick Prescott filter for the sales of item \( j_c \) on day \( t \) is obtained from minimizing (Hodrick and Prescott, 1993):

\[
\sum_{t=1}^{T} (ZSALES_{j,t} - HP_{j,t}) + \lambda \sum_{t=2}^{T-1} ((HP_{j,t+1} - HP_{j,t}) - (HP_{j,t} - HP_{j,t-1}))^2
\]

(5.6)
Sales decomposition using daily data for a single store

with

\[ ZSALES_{j,t} = \text{standardized sales } (ZSALES_{j,dw}) \text{ in week } w \text{ and day } d \text{ and indexed by } t, \ t=1,\ldots,D\times W, \]

\[ HP_{j,t} = \text{the value of the Hodrick Prescott filter for sales observation } t \text{ on item } j, \]

\[ \lambda = \text{parameter that determines which part of the data is removed.} \]

The filtered standardized sales are computed by subtracting the estimated value of the filter from the standardized sales (step 4):

\[ SALES_{j,t} = ZSALES_{j,t} - HP_{j,t} \]  

(5.7)

Re-indexing \( SALES_{j,t} \) using the relationship between the indices \( d, w, \) and \( t \) gives \( SALES_{j,dw} \). The parameter \( \lambda \) in (5.6) is called the smoothing parameter and is restricted to be positive. Higher values for \( \lambda \) lead to a smoother long-term component. In one extreme case \((\lambda = \infty)\) we obtain a filter that equals a linear trend. In the other extreme case \((\lambda = 0)\) the filter is equal to the observed sales.

Hodrick and Prescott (1993) propose different values for \( \lambda \) that depend on the data’s observation frequency (monthly, quarterly etc.). We use the suggested value for weekly data4. We tested the sensitivity for different values for \( \lambda \) following the recommendations in Hodrick and Prescott (1993). Changes appeared to be minor. The average significant (at \( p=0.05 \)) parameter estimate changes by 8 percent and the maximum change is 20 percent5.

We apply the Hodrick Prescott filter to both the sales and predictor variables (step 5). Not detrending the marketing instruments would lead to biased estimates because detrending removes (distracts) part of the instrument’s effect from the sales variable (see e.g. Abraham and Lodish, 1987, 1993). Consequently, the estimated instrument effects are based on the filtered sales (which are lower) and

---

4 The suggested value for the Hodrick Prescott filter is computed from the formula \( \lambda = (10 \times FREQ)^2 \) with \( FREQ \) frequency of observations defined as the number of observations per year. In our case the number of weeks is 53, and hence the value for \( \lambda = 280900 \). We use the suggestion for weekly data as we consider daily observations as repeated measurements from one week.

5 We also estimated the model without filtering. The estimates then change considerably.
Chapter 5

are thus underestimated. Thus, we have filtered price indices (FPI) and filtered feature (FFEAT) as predictor variables.

(iii) Other specification issues
Disaggregate sales
The standard approach of Van Heerde et al. (2003a) decomposes the promotion effect on own-brand sales into cross-brand-, cross-period-, and category expansion effects. This decomposition implicitly assumes that all relevant effects take place at the brand, cross-brand (all other brands together) and cross-period levels (all other periods together). We expect cross-brand and cross-category effects to occur at the item (SKU) level. This assumption is based on the observation that some items within and between categories are more similar than others (e.g. same brand or positioning). Therefore, we specify our models using item-level filtered standardized sales as the criterion variable.

Specifying models at the item level has two additional advantages. First, different (extended) decompositions can be computed directly (e.g. effects on different varieties). Second, we don’t have to estimate a set of models for different sets of predictor variables for each focal brand.

Leads and lags
Van Heerde et al. (2003a) include leads and lags among the predictor variables. The lead- and lagged variables accommodate the effects of promotional instruments in weeks surrounding the week of a promotion. We considered the inclusion of lead and lags. Modeling leads and lags for daily data is more complex than for weekly data as lagged variables for consecutive days are highly collinear. Additionally price variation is frequently so high (every second week) that including leads and lags makes no sense. Therefore we do not account for lead and lags.

6 Abraham and Lodish (1993, 1987) focus on baseline sales and therefore apply an iterative procedure to isolate the trend from the promotional sales. We focus on the instrument effects and therefore it suffices to filter both criterion- and predictor variables (see also Hodrick and Prescott, 1993).
7 This is different from Van Heerde et al. (2003b) who define focal brand dependent covariates.
5.3.2 Decompositions
We aim to decompose the price effects within- and across categories into effects on: (i) the item for which the price is changed (own-effect), (ii) other items of the same brand, and (iii) items of other brands. These effects can be both substitution- and complementary effects.

In terms of the variables we estimate the effect of the filtered price index ($FPI$) on sales defined as unit sales times average price ($SAP$). This effect cannot be directly estimated. We use the transformed variable $SALES$ (step 7 in Table 5.1). The consequence is that the parameter estimates have to be transformed back to represent the effect on $SAP$ (step 8 in Table 5.1). We show how to do this after the model specification. We first show the sales decomposition in terms of $SAP$. For this purpose we define two new indices that indicate whether items are of the same brand as item $j_c$:

- index $b_{c,j} : \text{items from category } c' \text{ of the same brand as } j_c \text{ (} j_c \text{ is excluded if } c' = c\text{)},$
- index, $r_{c,j} : \text{for the remaining items from category } c' \text{ that are from another brand as } j_c$.

For the decomposition we use the following criterion variables:

- sales of item $j_c$ ($SOI_{j,c}$),
- sales of other items of the same brand as item $j_c$ in category $c'$ ($SB_{j,c}$), to estimate within-brand effects (for the within-category decomposition we consider only category $c$),
- sales of items of other brands in category $c'$ ($SR_{j,c}$), to estimate cross-brand effects,
- total category sales for category $c$ ($SC_{j,c}$), to estimate the total effect on the category. We use this variable only for the within-category decomposition,
- total sales across all categories ($SAC_{j,c}$), to estimate the total effect across categories. We use this variable only for the cross-category decomposition.

These variables are defined as follows:

\[
SOI_{j,c} = SAP_{j,c}, \quad SB_{j,c} = \sum_{b_{c,j}} SAP_{j,c}, \quad SR_{j,c} = \sum_{r_{c,j}} SAP_{j,c},
\]

\[
SC_{c} = \sum_{j,c} SAP_{j,c}, \quad SAC_{c} = \sum_{c'} \sum_{j,c} SAP_{j,c}.
\]
Chapter 5

For the within-category decomposition we use the following identity which is true by construction:

\[ SC_{cdw} = SOI_{j,cdw} + SB_{j,cdw} + SR_{j,cdw} \tag{5.8} \]

This equation says that the total category sales equals the sum of the own-item sales, cross-item sales of items of the same brand, and sales of other items.

For the cross-category decomposition we use a similar identity:

\[ SAC_{dw} = SOI_{j,cdw} + SB_{j,cdw} + SR_{j,cdw} + \sum_{c' = 1}^{C} (SB_{j,c'dw} + SR_{j,c'dw}) \tag{5.9} \]

This equation says that the total sales across categories is equal to the sum of the own-item effect, the effect on items of the same brand within the category, effects on other items within the category, effects on items of the same brand in other categories and effects on items of other brands in other categories.

5.3.2.1 Within-category decomposition

We first estimate the effects of the filtered price index and filtered feature on filtered standardized sales (SALES). We specify an item-level linear regression model pooled across days for this purpose (step 7). We use these results to calculate the effects on the criterion variables for the decomposition (step 8). The model for item \( j_c \) is:

\[ SALES_{j,cdw} = C'_{j_c} + \sum_{j_c=1}^{J_c} (\beta_{j,cdw} FPI_{j,cdw} + \delta'_{j,cdw} FFEAT_{j,cdw}) + \epsilon_{j,cdw} \tag{5.10} \]

with \( \epsilon_{j,cdw} \) disturbance for day \( d \) in week \( w \). The parameter \( \beta_{j,cdw} \) in (5.10) is the effect of the filtered price index of item \( j_c \) on the filtered standardized sales of item \( j_c \). If \( j' = j_c \) we have the own effect (\( \beta_{j,cdw} \), expected to be negative), in all other cases we have cross-item effects.

We use the estimated parameters \( \hat{\beta}_{j,cdw} \) to calculate the decomposition. We show in Appendix 5B that multiplication of \( \hat{\beta}_{j,cdw} \) by the estimated standard deviation of sales (from 5.4) for a particular day \( \hat{\sigma}_{SAP,j,cdw} \) gives the day-specific effect on the filtered price index (FPI) on sales in units times average price (SAP). For our decomposition we report the average effect across days. We multiply this
Sales decomposition using daily data for a single store

effect by −1 to obtain the effect of a short-term price decrease instead of a price increase (step 8). Usually a price decrease leads to a sales increase. Thus, we have:

\[ \beta_{j,c}^* = -\hat{\sigma}_{SAP_j} \beta_{j,c} \]  

(5.11)

The coefficient \( \beta_{j,c}^* \) represents the average effect of the filtered price index on short term fluctuations in “sales” (SAP). We use these effects to calculate the effect on the criterion variables for the within-category decomposition: the own-item sales (SOI), sales of other items of brand \( j_c \) (SB), sales of other items (SR), and total category sales (SC).

We sum the item-level estimated effects (\( \beta_{j,c}^* \)) to obtain the price effect of item \( j_c \) on each component. Using (5.8) we have:

\[
\sum_{j_c=1}^{J} \beta_{j,c}^* = \beta_{j,c}^*_{SOI} + \sum_{b_j=1}^{B_j} \beta_{j,b_c}^* + \sum_{r_j=1}^{R_j} \beta_{j,r_c}^* \]

\( \Leftrightarrow \)

\[
\beta_{j,SC}^* = \beta_{j,SOI}^* + \beta_{j,SB}^* + \beta_{j,SR}^* \]  

(5.12)

Thus, the category effect \( \beta_{j,SC}^* \) is equal to the sum of the own effect \( \beta_{j,SOI}^* \) the effect on items of the same brand \( \beta_{j,SB}^* \) and the effect on items of other brands \( \beta_{j,SR}^* \).

For our application we use significant estimates to obtain reliable effects. We select estimates with a p-value less than 0.05. With these estimates we compute the relative sizes of category- and cross-item effects similar to Van Heerde et al. (2003a)\(^8\) (step 9). We test for the sensitivity in decompositions for the p-value used (0.01 to 0.10) in Section 5.5.5

5.3.2.2 Cross-category decomposition

A cross-category decomposition can be used to study the effect of an item’s filtered price on sales in other categories. We extend model (5.10) with instruments of items in other categories to estimate the cross-category

\(^8\) Note that in our application the cross effects have negative signs. In this way the results are comparable to the cross category decomposition.
decomposition effects needed for the decomposition (step 7). We use, as in (5.10), \(SALES_{j,d,w}\) as the criterion variable. The model for item \(j_c\) is:

\[
SALES_{j,d,w} = C_{j_c}^* + \sum_{c'=1}^{C} \sum_{j_c'=1}^{J_c} (\gamma^*_{j_c,j_c'} FPI_{j_c',d,w} + \delta^*_{j_c,j_c'} FFEAT_{j_c',d,w}) + \nu_{j,d,w} \tag{5.13}
\]

with

\[
\begin{align*}
\gamma^*_{j_c,j_c'} &= \text{parameter for the filtered price index of item } j'_{c'} \text{ on item } j_c, \\
\delta^*_{j_c,j_c'} &= \text{parameter for the filtered feature of item } j'_{c'} \text{ on item } j_c, \\
\nu_{j,d,w} &= \text{disturbance for day } d \text{ in week } w. 
\end{align*}
\]

The interpretation of the parameter \(\gamma^*_{j_c,j_c'}\) is the filtered price effect of item \(j'_{c'}\) on the sales of item \(j_c\). If \(c' = c\) we have the own-category effects (\(j'_{c'} = j_c\)) as in the within-category model. These effects include the own-item effects (\(j'_{c'} = j_c\)) and the cross-item effects (\(j'_{c'} \neq j_c\)). The parameters for which \(c' \neq c\) represent cross-category effects. We obtain \(\gamma^*_{j_c,j_c'}\), the effects on \(SAP\) (unit sales times average price), from a multiplication with minus the average standard deviation similar to (5.11) (step 8). We sum these parameters to obtain within- and cross-category effects on sales of other items of the same brand, items of other brands, and the total effect across categories. Using (5.9) we have:

\[
\sum_{c'=1}^{C} \sum_{j_c'=1}^{J_c} \gamma^*_{j_c,j_c'} = \gamma^*_{j_c,SOI} + \sum_{h=1}^{B_c} \gamma^*_{j_c,h_c} + \sum_{r=1}^{R_c} \gamma^*_{j_c,r_c} + \sum_{c'=1}^{C} \left( \sum_{b=1}^{B_{c'}} \gamma^*_{j_c,b_{c'}} + \sum_{r=1}^{R_{c'}} \gamma^*_{j_c,r_{c'}} \right) \\
\Leftrightarrow \quad \gamma^*_{j_c,SAC} = \gamma^*_{j_c,SOI} + \gamma^*_{j_c,SR_c} + \sum_{c'=1}^{C} \sum_{c \neq c'} \left( \gamma^*_{j_c,SR_{c'}} + \gamma^*_{j_c,SB_{c'}} \right) \tag{5.14}
\]

Thus, the effect on total sales across categories \(\gamma^*_{j_c,SAC}\) (expansion effect) is equal to the sum of the own-item effect \(\gamma^*_{j_c,SOI}\), effects on (i) other items of the same brand within the category \(\gamma^*_{j_c,SB_c}\), (ii) items of other brands in the same category \(\gamma^*_{j_c,SR_c}\), (iii) items of the same brand in other categories \(\gamma^*_{j_c,SB_{c'}}\), (iv) items of other brands in other categories \(\gamma^*_{j_c,SR_{c'}}\). We use the effects \(\gamma^*_{j_c,SAC}, \gamma^*_{j_c,SOI}, \gamma^*_{j_c,SB_c}, \gamma^*_{j_c,SR_c}, \gamma^*_{j_c,SB_{c'}}, \text{ and } \gamma^*_{j_c,SR_{c'}}\) to calculate the decomposition (step 9).

We accommodate substitution and complementary effects, by first summing all estimated effects with positive signs, and separately all effects with a negative sign. The positive effect is:
Sales decomposition using daily data for a single store

\[
\gamma^*_{j,\text{POS}} = \gamma^*_{j,\text{SOI}} + I_s(\gamma^*_{j,\text{SRi}})\gamma^*_{j,\text{SRi}} + I_s(\gamma^*_{j,\text{SRj}})\gamma^*_{j,\text{SRj}} + \sum_{c=1}^{C} (I_s(\gamma^*_{j,\text{SRi}})\gamma^*_{j,\text{SRi}} + I_s(\gamma^*_{j,\text{SRj}})\gamma^*_{j,\text{SRj}})
\]

(5.15) and the negative effect is:

\[
\gamma^*_{j,\text{NEG}} = (1 - I_s(\gamma^*_{j,\text{SRi}}))\gamma^*_{j,\text{SRi}} + (1 - I_s(\gamma^*_{j,\text{SRj}}))\gamma^*_{j,\text{SRj}} + \sum_{c=1}^{C} ((1 - I_s(\gamma^*_{j,\text{SRi}}))\gamma^*_{j,\text{SRi}} + (1 - I_s(\gamma^*_{j,\text{SRj}}))\gamma^*_{j,\text{SRj}})
\]

(5.16)

\(I_s()\) a function that takes value 1 if the argument is positive, and zero if not. Note that the negative effect has a negative sign. The sum of the positive and negative effects equals the expansion effect across categories \(\gamma^*_{j,\text{SAC}}\). The positive-effect has a similar interpretation as the own-item effect in the within category decomposition and is independent of the category definition. That is, excluding cross-category effects leads to a positive effect that equals the own-effect (assuming that the within-category effects on other items are negative).

We express all effects as a fraction of the positive effect. The fractions for the negative and total effects across categories reflect the relative sizes of the losses (negative effect) and the net-additional sales (total effect):

\[
\text{fraction negative effect: } \frac{\gamma^*_{j,\text{NEG}}}{\gamma^*_{j,\text{POS}}}
\]

\[
\text{fraction expansion effect: } \frac{\gamma^*_{j,\text{SAC}}}{\gamma^*_{j,\text{POS}}} = 1 + \frac{\gamma^*_{j,\text{NEG}}}{\gamma^*_{j,\text{POS}}}
\]

Expressing the components of the positive and negative effects in terms of the positive effect provides insight in the composition of these effects. We also calculate fractions for effects underlying the positive- and negative effects. The fractions related to the positive effect are:

\[
\text{fraction own effect: } \frac{\gamma^*_{j,\text{SOI}}}{\gamma^*_{j,\text{POS}}}
\]
Chapter 5

fraction positive effect on other items of the same brand within the category: 
\[
\frac{I_S(\gamma_{j,SB})\gamma_{j,SB}}{\gamma_{j,POS}}
\]

fraction positive effect on items of other brands within the category: 
\[
\frac{I_S(\gamma^*_{j,SR})\gamma^*_{j,SR}}{\gamma^*_{j,POS}}
\]

fraction positive effect on items of the same brand in category \(c':\) 
\[
\frac{I_S(\gamma^*_{j,SR})\gamma^*_{j,SR}}{\gamma^*_{j,POS}}
\]

fraction positive effect on items of other brands in category \(c':\) 
\[
\frac{I_S(\gamma^*_{j,SR})\gamma^*_{j,SR}}{\gamma^*_{j,POS}}
\]

The sum of these fractions is equal to 1. In a similar way we calculate fractions for the negative effect:

fraction negative effect on other items of the same brand within the category: 
\[
\frac{(1-I_S(\gamma_{j,SB})\gamma_{j,SB})}{\gamma_{j,POS}}
\]

fraction negative effect on items of other brands within the category: 
\[
\frac{(1-I_S(\gamma^*_{j,SR})\gamma^*_{j,SR})}{\gamma^*_{j,POS}}
\]

fraction negative effect on items of the same brand in category \(c':\) 
\[
\frac{(1-I_S(\gamma^*_{j,SR})\gamma^*_{j,SR})}{\gamma^*_{j,POS}}
\]

fraction negative effect on items of other brands in category \(c':\) 
\[
\frac{(1-I_S(\gamma^*_{j,SR})\gamma^*_{j,SR})}{\gamma^*_{j,POS}}
\]

The sum of these fractions equals the fraction negative effect.

86
5.4 Data

We apply the models to item level daily data of one Spanish hypermarket. We use one year of observations (311 days). We perform three separate studies on clusters of product categories for which we expect dependencies. The clusters are: (1) the sub-categories canned- and bottled beers, (2) the sub-categories concentrated- and non-concentrated dish detergents, (3) clothing detergents and the sub-categories concentrated and non-concentrated fabric softeners. We choose related categories based on the results of previous studies that show interdependencies primarily in related categories (see Section 5.2.1).

For each category we include items that have an annual market share of 5 percent or more and with sales data for more than 250 days. We provide descriptive statistics in Table 5.2. Table 5.2 shows per item the total annual turnover, the use of price changes and feature, the average price and the number of days with nonzero sales. The turnover is measured in million Pesetas. One million Pesetas is approximately €6000. Short term price variation is measured in two different ways. The average weekly price change compares the price of one day with the price on the same weekday in the previous week. We also report the percentage of days with a price change exceeding 5 percent. We define prices in terms of 1000 Pesetas.

In the beer categories a few items account for most of the sales. The selected items’ total share as a percentage of category sales (in terms of revenue) is 88 percent for the bottled beer category and 93 percent for the canned beer category. We see that the canned beers on average generate about three times as much revenues as bottled beers. Each of the four brands has one item in each category.

For both dish detergent sub-categories we have six items. The total share of the selected items is 90 percent for the concentrated dish detergent category and 83 percent for the non-concentrated category. The annual turnover in these categories is relatively small. There are two brands in the concentrated category: Fairy and Mistol. In the non-concentrated category we have three brands: Coral, Flota, and Mistol. Mistol is the only brand with items in both categories.

We have eight items for the concentrated fabric softener sub-category, five items for the non-concentrated sub-category and six items for the clothes detergents category. The selected items’ share is low compared to the other categories (63 percent for concentrated dish detergents, 58 percent for non-concentrated dish detergents, and 52 percent for clothes detergents). These low percentages result from the fact that there are many items in each of these
Chapter 5

categories. Sales (revenue) of the concentrated- and non-concentrated fabric softeners are low compared to clothes detergents. There are four brands with items in both fabric softener sub-categories. There are no clothes detergent brands with items in the fabric softener categories. Importantly, we see that price variation in the fabric softener and clothes detergents categories is small compared to the variation in prices in the beer and dish detergent categories.

5.5 Results
We explain first how we treat correlated covariates (Section 5.5.1). Then we show the within-category decomposition. We show item level results for the concentrated dish detergent’s category and average results for all other categories. In Section 5.5.3 we show the cross-category decompositions. For the cross-category decompositions we consider dependencies between pairs of categories. Due to the modes sample size, we cannot consider more than two categories at the same time. We show for the within category-decomposition, item-level results for the concentrated dish detergent category, and averages for all other categories. The within- and cross category decompositions are compared in Section 5.5.4. In Section 5.5.5 we test for the sensitivity of the p-value used to select effects for the decomposition.

5.5.1 Correlated covariates
For 11 groups of covariates correlations are too high to obtain good estimates. We show in Table 5.3 how we proceed based on the approach explained in Section 5.3.1.
## Table 5.2 Item Level Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Annual turnover in million PTAS</th>
<th>Percentage of weekly price changes exceeding 5%</th>
<th>Average weekly price change (in %)</th>
<th>No of days with feature</th>
<th>Average price in PTAS</th>
<th>No. of days with sales</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bottled beer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aguila</td>
<td>3.2</td>
<td>0.3%</td>
<td>0.5%</td>
<td>0</td>
<td>121</td>
<td>301</td>
</tr>
<tr>
<td>Riders</td>
<td>1.4</td>
<td>39.0%</td>
<td>4.8%</td>
<td>132</td>
<td>96</td>
<td>298</td>
</tr>
<tr>
<td>Biederland</td>
<td>0.8</td>
<td>42.7%</td>
<td>6.6%</td>
<td>121</td>
<td>95</td>
<td>270</td>
</tr>
<tr>
<td>San Miguel</td>
<td>2.0</td>
<td>11.1%</td>
<td>1.0%</td>
<td>6</td>
<td>119</td>
<td>299</td>
</tr>
<tr>
<td><strong>Canned beer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aguila</td>
<td>9.4</td>
<td>12.4%</td>
<td>1.1%</td>
<td>24</td>
<td>52</td>
<td>301</td>
</tr>
<tr>
<td>Riders</td>
<td>5.5</td>
<td>40.5%</td>
<td>6.2%</td>
<td>125</td>
<td>38</td>
<td>301</td>
</tr>
<tr>
<td>Biederland</td>
<td>4.6</td>
<td>45.2%</td>
<td>8.1%</td>
<td>124</td>
<td>38</td>
<td>288</td>
</tr>
<tr>
<td>San Miguel</td>
<td>2.4</td>
<td>24.0%</td>
<td>2.9%</td>
<td>12</td>
<td>54</td>
<td>291</td>
</tr>
<tr>
<td><strong>Concentrated dish detergents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairy 0.75L Normal</td>
<td>0.5</td>
<td>25.0%</td>
<td>2.7%</td>
<td>18</td>
<td>251</td>
<td>297</td>
</tr>
<tr>
<td>Fairy 1L Normal</td>
<td>0.9</td>
<td>12.9%</td>
<td>1.9%</td>
<td>46</td>
<td>286</td>
<td>301</td>
</tr>
<tr>
<td>Fairy 0.75L Antibact</td>
<td>0.7</td>
<td>24.5%</td>
<td>4.0%</td>
<td>48</td>
<td>281</td>
<td>296</td>
</tr>
<tr>
<td>Mistol 0.75L Normal</td>
<td>0.3</td>
<td>19.3%</td>
<td>2.8%</td>
<td>13</td>
<td>221</td>
<td>291</td>
</tr>
<tr>
<td>Mistol 0.75L Antibact</td>
<td>0.5</td>
<td>21.9%</td>
<td>2.4%</td>
<td>17</td>
<td>221</td>
<td>299</td>
</tr>
<tr>
<td>Mistol 0.75L pH5.5</td>
<td>0.3</td>
<td>24.2%</td>
<td>3.0%</td>
<td>13</td>
<td>224</td>
<td>291</td>
</tr>
<tr>
<td><strong>Non-concentrated dish detergents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coral 1L Normal</td>
<td>0.3</td>
<td>15.7%</td>
<td>2.9%</td>
<td>0</td>
<td>96</td>
<td>282</td>
</tr>
<tr>
<td>Coral 1.5L Normal</td>
<td>0.7</td>
<td>15.7%</td>
<td>1.9%</td>
<td>48</td>
<td>118</td>
<td>301</td>
</tr>
<tr>
<td>Flota 1.5L Normal</td>
<td>0.3</td>
<td>19.3%</td>
<td>2.9%</td>
<td>49</td>
<td>104</td>
<td>300</td>
</tr>
<tr>
<td>Mistol 1L Normal</td>
<td>0.9</td>
<td>23.2%</td>
<td>3.5%</td>
<td>12</td>
<td>102</td>
<td>300</td>
</tr>
<tr>
<td>Mistol 1.5L Normal</td>
<td>0.4</td>
<td>15.9%</td>
<td>2.3%</td>
<td>12</td>
<td>148</td>
<td>295</td>
</tr>
<tr>
<td>Mistol 1L Lemon</td>
<td>0.3</td>
<td>16.4%</td>
<td>2.1%</td>
<td>12</td>
<td>100</td>
<td>301</td>
</tr>
<tr>
<td><strong>Concentrated fabric softeners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flor 1L Normal</td>
<td>0.9</td>
<td>0.7%</td>
<td>0.6%</td>
<td>0</td>
<td>283</td>
<td>301</td>
</tr>
<tr>
<td>Flor 0.5L Blue</td>
<td>0.4</td>
<td>10.1%</td>
<td>1.2%</td>
<td>6</td>
<td>154</td>
<td>298</td>
</tr>
<tr>
<td>Flor 0.5L White</td>
<td>0.3</td>
<td>12.0%</td>
<td>0.8%</td>
<td>6</td>
<td>154</td>
<td>266</td>
</tr>
<tr>
<td>San 0.5L Normal</td>
<td>0.3</td>
<td>5.3%</td>
<td>0.9%</td>
<td>0</td>
<td>108</td>
<td>300</td>
</tr>
<tr>
<td>San 0.5L Wildflowers</td>
<td>0.3</td>
<td>5.8%</td>
<td>0.9%</td>
<td>0</td>
<td>108</td>
<td>293</td>
</tr>
<tr>
<td>Lenor 1L Normal</td>
<td>0.3</td>
<td>4.0%</td>
<td>0.3%</td>
<td>0</td>
<td>266</td>
<td>276</td>
</tr>
<tr>
<td>Quanto 0.5L Normal</td>
<td>0.4</td>
<td>10.4%</td>
<td>1.2%</td>
<td>0</td>
<td>134</td>
<td>299</td>
</tr>
<tr>
<td>Kel 0.75L Normal</td>
<td>0.3</td>
<td>5.0%</td>
<td>0.9%</td>
<td>0</td>
<td>93</td>
<td>300</td>
</tr>
<tr>
<td><strong>Non concentrated fabric softeners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flor 3.5L Blue</td>
<td>1.1</td>
<td>2.5%</td>
<td>0.1%</td>
<td>96</td>
<td>373</td>
<td>301</td>
</tr>
<tr>
<td>Lenor 4L Normal</td>
<td>0.9</td>
<td>2.5%</td>
<td>0.2%</td>
<td>62</td>
<td>351</td>
<td>285</td>
</tr>
<tr>
<td>Quanto 4L Normal</td>
<td>1.2</td>
<td>5.9%</td>
<td>0.5%</td>
<td>61</td>
<td>319</td>
<td>297</td>
</tr>
<tr>
<td>Kel 4L Normal</td>
<td>0.4</td>
<td>7.9%</td>
<td>0.4%</td>
<td>96</td>
<td>222</td>
<td>288</td>
</tr>
<tr>
<td>Vernel 3L Blue</td>
<td>0.8</td>
<td>1.6%</td>
<td>1.0%</td>
<td>65</td>
<td>363</td>
<td>293</td>
</tr>
<tr>
<td><strong>Clothes detergents</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ariel 2Kg</td>
<td>4.3</td>
<td>1.4%</td>
<td>0.9%</td>
<td>96</td>
<td>676</td>
<td>293</td>
</tr>
<tr>
<td>Ariel 2.9Kg</td>
<td>3.6</td>
<td>11.1%</td>
<td>1.2%</td>
<td>36</td>
<td>991</td>
<td>298</td>
</tr>
<tr>
<td>Wipp Progres 2Kg</td>
<td>2.3</td>
<td>2.0%</td>
<td>1.2%</td>
<td>71</td>
<td>657</td>
<td>299</td>
</tr>
<tr>
<td>Colon 2Kg</td>
<td>2.9</td>
<td>8.6%</td>
<td>1.6%</td>
<td>72</td>
<td>607</td>
<td>301</td>
</tr>
<tr>
<td>Elena 2Kg</td>
<td>2.3</td>
<td>9.2%</td>
<td>1.9%</td>
<td>90</td>
<td>480</td>
<td>294</td>
</tr>
<tr>
<td>Micolor 2Kg</td>
<td>3.0</td>
<td>0.0%</td>
<td>0.8%</td>
<td>24</td>
<td>676</td>
<td>301</td>
</tr>
</tbody>
</table>
Chapter 5

For six groups of variables a price decrease is almost always accompanied by feature. The feature variable is excluded in each of these cases so that the price effect really captures feature-supported effects. For one group of items we find that the store uses feature for two different items simultaneously (Mistol 0.75L Normal and Mistol 0.75L pH 5.5). In this case we include a common feature variable. We find that prices are frequently used simultaneously for four groups of items of the same brand. We construct new variables for these items, a common price (if prices are the same) and we construct separate price variables for each item if prices are not the same.

5.5.2 Within-category decomposition

We show the within-category decomposition for the concentrated dish detergents in Table 5.4. The second column of this table shows the own effects. The third column shows the effect on other items of the same brand as a percentage of the own item effect. The fourth column shows the effect on items of other brands. The last column shows the percentage category expansion effect. For example we observe for Fairy 0.75 Antibacterial that the own effect is 7.2. This implies an average sales increase of 720 pesetas (about €4) for a 10 percent price decrease\(^9\). 60 percent of this sales increase comes from sales of other items of Fairy, 11 percent is drawn from sales of items belonging to other brands. Hence, the total primary demand effect is 29 percent (100 percent – 60 percent – 11 percent) of the own effect in terms of revenue. If we only subtract the effect on the manufacturer’s items, we obtain an expansion effect of 40 percent (100 percent – 60 percent).

For the other items in the category we find own effects varying between 2.1 and 11.8. Price changes for each of the other items of Fairy affect the sales of other items of Fairy negatively. This means that the revenue for the manufacturer and for the retailer are negatively affected by these cross effects. For Mistol’s items we find an average own effect of 5.3, an effect on items of the same brand of –29 percent and –2 percent on items of other brands, and hence we find an expansion effect of 70 percent.

---

\(^9\) We use prices in terms of 1000 Pesetas for modeling. This implies that the coefficient represents the price effect in terms of 1000 Pesetas for a 100 percent decrease.

90
Table 5.3 Multicollinearity: problems and solutions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Problem</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riders 1L Price</td>
<td>High correlation (-0.998)</td>
<td>We exclude Riders 1L feature</td>
</tr>
<tr>
<td>Riders 1L Feature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riders 0.33L Price</td>
<td>High correlation (-0.976)</td>
<td>We exclude Riders 0.33L feature</td>
</tr>
<tr>
<td>Riders 0.33L Feature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mistol 0.75L Normal feature</td>
<td>Always used simultaneously</td>
<td>We use a combined feature variable</td>
</tr>
<tr>
<td>Mistol 0.75L pH 5.5 feature</td>
<td>(correlation is 1)</td>
<td></td>
</tr>
<tr>
<td>Mistol 0.75L Normal price</td>
<td>Strong correlations</td>
<td>We construct four variables, common price three items, unique price for each item</td>
</tr>
<tr>
<td>Mistol 0.75L Antibact price</td>
<td>(Normal-Antibact: 0.672)</td>
<td></td>
</tr>
<tr>
<td>Mistol 0.75L pH5.5 price</td>
<td>Normal-pH: 0.650</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Antibact-pH: 0.708</td>
<td></td>
</tr>
<tr>
<td>Mistol 1L Normal price, Mistol 1L Lemon price</td>
<td>Strong correlation (0.752)</td>
<td>We construct three variables, common price two items, unique prices for each item</td>
</tr>
<tr>
<td>San 0.5L Normal price</td>
<td>Always used simultaneously</td>
<td>We use a common price variable</td>
</tr>
<tr>
<td>San 0.5L Wildflowers price</td>
<td>(correlation is 1)</td>
<td></td>
</tr>
<tr>
<td>Flor 0.5L Blue price</td>
<td>Always used simultaneously</td>
<td>We use a common price variable</td>
</tr>
<tr>
<td>Flor 0.5L White price</td>
<td>(Correlation is 1)</td>
<td></td>
</tr>
<tr>
<td>Vernel 3L Normal feature</td>
<td>High correlation (-0.934)</td>
<td>We exclude Vernel 3L feature</td>
</tr>
<tr>
<td>Vernel 3L Normal price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kel 4L Normal feature</td>
<td>High correlation (-0.831)</td>
<td>We exclude Kel 4L feature</td>
</tr>
<tr>
<td>Kel 4L Normal price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Micolor 2Kg feature</td>
<td>High correlation (-0.895)</td>
<td>We exclude Micolor Solido 4Kg feature</td>
</tr>
<tr>
<td>Micolor 2Kg price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elena 2Kg feature</td>
<td>High correlation (-0.921)</td>
<td>We exclude Elena 2Kg feature</td>
</tr>
<tr>
<td>Elena 2Kg price</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aCorrelations are calculated between price- and feature variables after deseasonalizing and detrending.
Table 5.4 Within category decompositions for concentrated dish detergents

<table>
<thead>
<tr>
<th></th>
<th>Own</th>
<th>Within category on other items of the same brand</th>
<th>Within category on items of other brands</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Fairy 0.75L Normal, Price</td>
<td>7.3</td>
<td>-45%</td>
<td>0%</td>
<td>55%</td>
</tr>
<tr>
<td>Price Fairy 1L Normal, Price</td>
<td>11.8</td>
<td>-44%</td>
<td>0%</td>
<td>56%</td>
</tr>
<tr>
<td>Price Fairy 0.75L Antibact, Price</td>
<td>7.2</td>
<td>-60%</td>
<td>-11%</td>
<td>29%</td>
</tr>
<tr>
<td>Mistol 0.75L Normal/Antibact/pH 5.5, Common price</td>
<td>2.4</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Mistol 0.75L Normal, Unique price</td>
<td>2.1</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Mistol 0.75L pH 5.5, Unique price</td>
<td>3.3</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Average</td>
<td>5.3</td>
<td>-29%</td>
<td>-2%</td>
<td>70%</td>
</tr>
</tbody>
</table>

Table 5.5 Within category decompositions, averages per category

<table>
<thead>
<tr>
<th></th>
<th>Own</th>
<th>Within category on other items of the same brand</th>
<th>Within category on items of other brands</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottled beer</td>
<td>52.6</td>
<td>0%</td>
<td>-18%</td>
<td>82%</td>
</tr>
<tr>
<td>Canned beer</td>
<td>103.6</td>
<td>0%</td>
<td>-20%</td>
<td>80%</td>
</tr>
<tr>
<td>Concentrated dish detergents</td>
<td>5.3</td>
<td>-29%</td>
<td>-2%</td>
<td>70%</td>
</tr>
<tr>
<td>Non-concentrated dish detergents</td>
<td>10.2</td>
<td>-19%</td>
<td>-14%</td>
<td>66%</td>
</tr>
<tr>
<td>Concentrated fabric softeners</td>
<td>5.0</td>
<td>0%</td>
<td>-48%</td>
<td>52%</td>
</tr>
<tr>
<td>Non-concentrated fabric softeners</td>
<td>10.2</td>
<td>0%</td>
<td>-49%</td>
<td>51%</td>
</tr>
<tr>
<td>Clothes detergents</td>
<td>31.7</td>
<td>0%</td>
<td>-81%</td>
<td>19%</td>
</tr>
<tr>
<td>Average</td>
<td>31.2</td>
<td>-7%</td>
<td>-33%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Table 5.5 shows the average decomposition per category. The average own-item effects are large for beer, smaller for clothes detergents and smallest for dish detergents and fabric softeners. The differences can be partially explained by the difference in average item turnover between categories (larger effects for items with a higher turnover).

We find effects on other items of the same brand only in the concentrated and non-concentrated dish detergent categories. Note that not all categories have multiple items of the same brand. For all categories, we find effects on other items
Sales decomposition using daily data for a single store

in the category. The effects range from –2 percent for concentrated dish detergents to –81 percent for clothes detergents. The expansion effects (100 percent minus both cross effects) vary between 19 percent (clothes detergents) and 82 percent (bottled beer).

The average own effect across all categories is 31.2. On average 7 percent of this effect is lost to items of the same brand and 33 percent is lost to items of other brands. This leads to an average cross item effect of 40 percent (33 percent + 7 percent) and an expansion effect of 60 percent. The average cross item effect is close to the 32 percent average found by Van Heerde et al. (2003b). If we only consider the effects on the items for the manufacturer, we obtain larger expansion effects. The average expansion effect if we subtract only the effect on manufacturers’ own items is 93 percent (100 percent –7 percent).

5.5.3 Cross-category decomposition
We show the results for the cross-category decompositions for the concentrated dish detergents with respect to non-concentrated dish detergents in Table 5.6. The columns of Table 5.6 are divided in three parts. The first part summarizes the decomposition into positive-, negative- and expansion effects. The next five columns decompose the positive effect into the (i) own effect, (ii) effect on other items of the same brand within the category, (iii) effect on items of other brands within the category, (iv) effect on items of the same brand in another category, (v) effect on items of other brands in other categories. The last four columns show a decomposition of the negative effects.

For example we observe for Fairy 0.75L Antibacterial a positive effect of 8.1. This implies an average sales increase of 810 pesetas (about €5) for a 10 percent price decrease. The store loses about half of this increase (49 percent; the negative effect) to other items. Hence, the net sales increase or expansion effect is (51 percent) of the positive effect. The next columns (columns 5 - 9) show the decomposition of the positive effect. We see that the own-effect entirely accounts for the positive effect and that there are no positive effects on other items within- and across categories. The last two columns decompose the negative effects. The 49 percent sales loss (negative effect) is due to a 29 percent loss to items of the same brand and a 20 percent loss to items of other brands. If we consider the

---

10 The estimated own effect of 8.1 is different from the estimate of 7.2 in Table 5.4 due to differences in model specifications. We expand on this issue in Section 5.5.4.
Chapter 5

effect on the manufacturer’s items only, we find an expansion effect of 71 percent (100 percent – 29 percent).

For the other items in the category we have positive effects varying between 1.9 and 12.3. The average effect is 6.0. The average negative effect is 51 percent of the positive effect. For two items we find no negative effects and we find one negative effect that exceeds the positive effect (166 percent for the common price of Mistol 0.75L) though the confidence interval includes 100 percent. Note that contrary to the within category decomposition we find no within category effect on items of the same brand for Fairy 0.75L Normal. This is due to the fact that the corresponding parameter estimate is no longer significant.

Table 5.7 shows the average decomposition for each category pair considered. The average positive effects vary between 5.1 and 157.7. The effects are larger for categories with higher average sales (revenues) per item. We see that the average own effect for pairs of categories for which one category is the same varies somewhat (e.g. non concentrated fabric softener on concentrated fabric softener and detergents). This can be explained from the fact that we estimate different models with different predictor variables. This leads in general to different results. The same phenomenon also affects the within category effects. This problem would not occur if we consider all categories simultaneously instead of category pairs. (We do not do this as this results in models with too many variables.)

The average negative effects vary from –14 percent to –56 percent and hence the expansion effect varies from 44 percent to 86 percent (100 percent – the negative effect). The expansion effect is largest for beer and fabric softeners. The average negative effect across all category pairs is 36 percent and hence the expansion effect is 64 percent.

The decomposition of the positive effect shows, as expected, that the own effect accounts for the largest part of the positive effect (88 percent on average across all category pairs). There is no evidence for positive cross effects on items of the same brand within the category. We find a positive effect on items of the other brands within the category in four categories. The average of this effect across all category pairs is 3 percent. Positive cross category effects on items of the same brand are virtually nonexistent (1 percent on average across all category pairs). We find stronger positive cross category effects on items of other brands (8 percent on average).
### Table 5.6 Cross category decomposition concentrated dish detergents

<table>
<thead>
<tr>
<th>Product Description</th>
<th>Positive</th>
<th>Negative</th>
<th>Expansion</th>
<th>Within Own Brand</th>
<th>Within Other Brands</th>
<th>Cross Own Brand</th>
<th>Cross Other Brands</th>
<th>Decomposition positive effect</th>
<th>Decomposition negative effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairy 0.75L Normal, Price</td>
<td>8.0</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fairy 1L Normal, Price</td>
<td>12.3</td>
<td>-41%</td>
<td>59%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td>-41%</td>
<td></td>
</tr>
<tr>
<td>Fairy 0.75L Antibact, Price</td>
<td>8.1</td>
<td>-49%</td>
<td>51%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td>-29%</td>
<td>-20%</td>
</tr>
<tr>
<td>Mistol 0.75L Normal/Antibact/pH 5.5, Common Price</td>
<td>1.9</td>
<td>-166%</td>
<td>-66%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td>-166%</td>
<td></td>
</tr>
<tr>
<td>Mistol 0.75L Normal, Unique Price</td>
<td>2.5</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mistol 0.75L pH 5.5, Unique Price</td>
<td>3.5</td>
<td>-51%</td>
<td>49%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td>-51%</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>6.0</td>
<td>-51%</td>
<td>49%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>-20%</td>
</tr>
</tbody>
</table>
### Table 5.7 Cross category decomposition, averages across category pairs

<table>
<thead>
<tr>
<th>Category Pair</th>
<th>Decomposition positive effect</th>
<th>Decomposition negative effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within category</td>
<td>Cross category</td>
</tr>
<tr>
<td></td>
<td>Own Same brand</td>
<td>Other brands</td>
</tr>
<tr>
<td>Bottled beer on canned beer</td>
<td>Positive 53.2</td>
<td>Negative -16%</td>
</tr>
<tr>
<td>Canned beer on bottled beer</td>
<td>Positive 157.7</td>
<td>Negative -20%</td>
</tr>
<tr>
<td>Concentrated dish detergents on</td>
<td>Positive 6.0</td>
<td>Negative -51%</td>
</tr>
<tr>
<td>Non-concentrated dish detergents on</td>
<td>Positive 8.9</td>
<td>Negative -47%</td>
</tr>
<tr>
<td>Concentrated fabric softeners on</td>
<td>Positive 8.2</td>
<td>Negative -17%</td>
</tr>
<tr>
<td>Non-concentrated fabric softeners on</td>
<td>Positive 10.2</td>
<td>Negative -56%</td>
</tr>
<tr>
<td>Concentrated fabric softener on</td>
<td>Positive 17.8</td>
<td>Negative -25%</td>
</tr>
<tr>
<td>Detergents on non-concentrated fabric</td>
<td>Positive 35.2</td>
<td>Negative -50%</td>
</tr>
<tr>
<td>softener</td>
<td>Positive 5.1</td>
<td>Negative -14%</td>
</tr>
<tr>
<td>Average</td>
<td>Positive 33.1</td>
<td>Negative -36%</td>
</tr>
</tbody>
</table>
The decomposition of the negative effects shows that the within category effect on other brands is larger than the effects on other items of the same brand. We note that the number of brands with multiple items is limited. The average effect on items of other brands across all category pairs is −26 percent. The category averages vary between −2 percent (non-concentrated fabric softener on detergents) and −65 percent (detergents on concentrated fabric softener). Within category effects on items of the same brand only occur within the concentrated dish detergents categories. The average across all categories is −3 percent. We find no negative cross category effects on items of the same brand, but we do find negative cross category effects on items of other brands for seven out of eleven category pairs. The average negative effect on items of other brands is −10 percent across all category pairs.

5.5.4 Comparison of the within- and cross-category decompositions

One way to assess the impact of considering cross category effects is to compare the within- and cross category decompositions. The problem with this comparison is that the results are affected by (i) the change in estimates due to inclusion of additional predictor variables (e.g. the own effects change) and (ii) the cross category effects. The first factor is inherent to variable addition unless the current and added variables are uncorrelated. The second factor is the effect of interest. We argue that the effect of considering cross category effects can be best assessed from the cross category decomposition only. That is, we consider the magnitude of the cross category effects.

Using this approach we see that considering cross category effects does matter. We find an average positive cross category effect of 9 percent (1 percent + 8 percent, effects on items of the same brand and other brands together) and an average negative cross category effect of −10 percent. Importantly, these cross effects are mostly effects on other brands. We recognize that this depends on the categories selected. The cross category effects, on average, are about one third of the size of the cross effects within the category. This is based on category pairs. Therefore we expect these effects to be stronger if more than two categories are considered simultaneously.
5.5.5 Sensitivity for the p-value used to obtain reliable decomposition results

The decompositions in the previous sections are based on estimates for which the two-sided p-value is below 0.05. In this section, we test the decomposition’s sensitivity for alternative p-values by comparing the outcomes for p=0.01 and p=0.10. We proceed as follows. First, we calculate for both p-values the within- and cross-category decompositions at the item level. Second, we determine the absolute difference with the decompositions based on p=0.05. For consistency and comparability, we do this only for items with a significant own effect at p=0.01. Finally, we calculate the average of both differences per category. In this way we obtain tables similar to Tables 5.5 and 5.7 with sensitivity results per cell. For example the sensitivity of the fraction own effect for a category (SENS_{FOE}) is obtained as follows:

$$SENS_{FOE} = \frac{|FOE_{p=0.10} - FOE_{p=0.05}| + |FOE_{p=0.05} - FOE_{p=0.01}|}{2}$$ (5.17)

with $FOE_{p=X}$ the average fraction own effect for a category if $p = X$. For the own effect in the within category decomposition and the positive effect in the cross category decomposition we also calculate the relative change of the results. The average relative changes are indicated between brackets. This is not necessary for the other effects as these are expressed as a percentage.

Table 5.8 shows the sensitivity for the within category decomposition. On average there is very little sensitivity. We see that the sensitivity varies across categories. For example we find for canned beer that the results are very stable. Only the effect on items of other brands and the expansion effect change by 3 percent. Similarly, the results for bottled beers, concentrated dish detergent and non-concentrated fabric softeners are stable. The results for the other categories are less stable.

The difference in stability across categories is potentially attributable to the difference in price changes across categories. Less variation in prices will lead to less precisely estimated effects and hence to higher sensitivity for the p-value used. Table 5.2 (second column) shows that for beer and dish detergents the percentage of weekly price changes (exceeding 5 percent) is typically between 10 and 45 percent. We find stable results for three of these four categories (the exception is the non-concentrated dish detergent category). For the other categories (clothes detergents and fabric softeners) we find that the percentage of...
changes is between 2 and 10 percent. For two of these categories we find that the results are less stable (the exception is the clothes detergent category).

For the average across categories we see that the results are fairly stable. The within-category effect on items of the same brand changes on average by 3 percent, the within-category effect on items of other brands changes by 10 percent, and the expansion effect changes by 13 percent. These changes are about 1/3 of the values in Table 5.5.

Table 5.8 Within category decompositions, sensitivity analysis for changes in p-value (0.01 and 0.10)

<table>
<thead>
<tr>
<th>Category</th>
<th>Own (%)</th>
<th>Within category on other items of the same brand</th>
<th>Within category on items of other brands</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottled beer</td>
<td>0.0 (0%)</td>
<td>0%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>Canned beer</td>
<td>0.0 (0%)</td>
<td>0%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Concentrated dish detergents</td>
<td>0.3 (1%)</td>
<td>0%</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>Non-concentrated dish detergents</td>
<td>0.0 (0%)</td>
<td>18%</td>
<td>14%</td>
<td>30%</td>
</tr>
<tr>
<td>Concentrated fabric softeners</td>
<td>0.6 (12%)</td>
<td>0%</td>
<td>26%</td>
<td>26%</td>
</tr>
<tr>
<td>Non-concentrated fabric softeners</td>
<td>0.0 (0%)</td>
<td>0%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Clothes detergents</td>
<td>0.0 (0%)</td>
<td>0%</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>Average</td>
<td>0.1 (2%)</td>
<td>3%</td>
<td>10%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 5.9 shows the sensitivity results for the cross-category decomposition. The average decomposition across all category pairs is fairly stable. The negative and expansion effects on average change by 18 percent. Important findings are that (1) the positive cross category effect changes on average by 5 percent (more than half of the effect found at p=0.05; 8 percent), (2) the negative effect on items of other brands within the category changes on average by 15 percent (about half of the effect found at p=0.05; −26 percent) (3), the effect on items of other brands in other categories changes by 3 percent (about 1/3 of the effect found at p=0.05; −10 percent).

For beer and dish detergents the results are quite stable. None of the changes exceeds 10 percent. For clothes detergents and fabric softeners the results are less stable. We find at least one major change for each of these pairs. We explain the
Chapter 5

difference in stability again by differences in price variation. That is, the results
for categories with more price variation are stable and results for categories with
less price variation are less stable.

We conclude that sensitivity for the p-value varies across categories. The
amount of price variation seems to be critical.
Table 5.9 Cross category decomposition, sensitivity analysis for changes in p-value (0.01 and 0.10)

<table>
<thead>
<tr>
<th></th>
<th>Decomposition positive effect</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within category</td>
<td>Cross category</td>
<td>Within category</td>
<td>Cross category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>Expansion</td>
<td>Within brand</td>
<td>Other brands</td>
<td>Same brand</td>
<td>Other brands</td>
<td>Same brand</td>
<td>Other brands</td>
<td>Same brand</td>
<td>Other brands</td>
</tr>
<tr>
<td>Bottled beer on canned beer</td>
<td>0.0 (0%)</td>
<td>3%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Canned beer on bottled beer</td>
<td>10.1 (3%)</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>3%</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Concentrated dish detergents on non-concentrated dish detergents</td>
<td>0.2 (2%)</td>
<td>10%</td>
<td>10%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>9%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>Non-concentrated dish detergents on concentrated dish detergents</td>
<td>0.0 (0%)</td>
<td>10%</td>
<td>10%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>0%</td>
<td>3%</td>
</tr>
<tr>
<td>Concentrated fabric softeners on non concentrated fabric softeners</td>
<td>2.6 (30%)</td>
<td>7%</td>
<td>7%</td>
<td>19%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>19%</td>
<td>0%</td>
<td>7%</td>
<td>0%</td>
</tr>
<tr>
<td>Non concentrated fabric softeners on concentrated fabric softeners</td>
<td>0.8 (6%)</td>
<td>20%</td>
<td>20%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>13%</td>
<td>3%</td>
</tr>
<tr>
<td>Non concentrated fabric softeners on detergents</td>
<td>4.5 (15%)</td>
<td>37%</td>
<td>37%</td>
<td>10%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>8%</td>
<td>0%</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>Detergents on non concentrated fabric softener</td>
<td>3.9 (7%)</td>
<td>30%</td>
<td>30%</td>
<td>7%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>0%</td>
<td>30%</td>
<td>0%</td>
</tr>
<tr>
<td>Concentrated fabric softener on detergents</td>
<td>2.5 (75%)</td>
<td>31%</td>
<td>31%</td>
<td>4%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
<td>12%</td>
<td>0%</td>
<td>47%</td>
<td>0%</td>
</tr>
<tr>
<td>Detergents on concentrated fabric softeners</td>
<td>1.5 (4%)</td>
<td>30%</td>
<td>30%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>3%</td>
<td>0%</td>
<td>28%</td>
<td>0%</td>
</tr>
<tr>
<td>Average</td>
<td>2.7 (14%)</td>
<td>18%</td>
<td>18%</td>
<td>5%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>5%</td>
<td>1%</td>
<td>15%</td>
<td>0%</td>
</tr>
</tbody>
</table>
5.6 Conclusions
In this chapter we develop a model to determine store-specific instrument effects using daily data. One benefit of using daily data is an increase in the number of observations. We use this model to decompose price promotion effects within- and across categories. Available decomposition methods based on elasticities or unit sales consider within-category effects only. Additionally we separate effects into effects on items of the same and other brands. We show that the omission of cross-category effects may lead to either over- or under estimation of the price promotion’s effect.

The fact that cross-category effects can be both positive and negative requires an approach different from the one used by Van Heerde et al. (2003a). For the cross category decomposition we separate the own- , cross-item- and cross category effects into sales increases and sales losses and sum the positive- and negative effects. The difference between the positive and the negative effect is the expansion effect. Next, we express the relative contribution of all different effects as fractions of the positive effect. We demonstrate the incremental value of the cross category decomposition over the within category decomposition developed by Van Heerde et al. (2003a).

The average cross-category decomposition across pairs of categories shows that 88 percent of the positive effect is due to the own effect. We find no positive within-category effects on items of the same brand, but we find an average positive within-category effect of 3 percent on items of other brands. The average positive cross effect on items of the same brand in another category is small (1 percent), while the average effect on items of other brands is larger (8 percent). The decomposition of the negative effect shows that the negative effect on items of other brands within the category is the strongest (-26 percent on average). The effect on items of the same brand within the category is considerably smaller (-3 percent on average). We find no negative cross category effects on items of the same brand, but we find an average effect of –10 percent on items of other brands.

Based on these results we conclude that considering cross-category effects has a serious impact on the evaluation of promotions. The positive and negative effects across categories are both about 1/3 of the within-category effects. We note that our findings are based upon pairs of categories. Larger cross-category effects (in total) are expected to be found if more categories are considered simultaneously.
Sales decomposition using daily data for a single store

We tested for the sensitivity of the p-value used to include effects in the decomposition. Our results suggest that the sensitivity varies across categories and that the sensitivity is higher for categories with less price variation. Thus, we conclude that daily data provide opportunities if the instruments vary sufficiently.

One limitation is that we consider only pairs of categories because the number of observations is too low to include all possible predictor variables. Future studies may consider alternative methods to solve this problem. Difficulties include: (i) if stepwise regression is used, the collinearity between predictor variables affects the outcomes of the stepwise procedure, and (ii) it is difficult to know which categories to include simultaneously, especially if the odds of finding cross-category effects are low.
Appendix 5A Price definition

Suppose one brand has two items $A$ and $B$ with prices $P_{At}$ and $P_{Bt}$ and sales $S_{At}$ and $S_{Bt}$. The prices of $A$ and $B$ are perfectly correlated during period $t = 1, \ldots, T$. In this case we cannot estimate the following linear models.

\begin{align*}
S_{At} &= c_A + \alpha_A P_{At} + \alpha_B P_{Bt} \\
S_{Bt} &= c_B + \beta_A P_{At} + \beta_B P_{Bt} \\
\end{align*}

(5A.1)

The solution is to leave out one of the price variables. The remaining price is defined as $P_{Ct}$ (a common price) and we estimate the following relations.

\begin{align*}
S_{At} &= c'_A + \alpha_C P_{Ct} \\
S_{Bt} &= c'_B + \beta_C P_{Ct} \\
\end{align*}

(5A.2)

Suppose now that we have a limited number of additional observations $(T + 1, \ldots, T^*)$ for which $P_{At}$ and $P_{Bt}$ differ. It is possible to estimate (5A.1) but the variance of the parameter estimates in (5A.1) is high because there is limited unique variation in the prices of $A$ and $B$ (multicollinearity). One might argue that from a managerial point of view it is appropriate to use a common price since the prices tend to move together.

In (5A.1) the common effect (the effect of simultaneous movement of prices) equals the sum of the own effects ($\alpha_A$, $\beta_B$) and the cross effects ($\alpha_B$, $\beta_A$). This, however, is based upon non-realistic assumptions:

- The cross effects for a common price change are the same as the cross effects for separate price changes. This assumption is not realistic as we expect no cross-effects between items that change the price at the same moment.
- The own-effect of a common price change is supposed to be equal to the sum of the own-effects of separate price changes. This assumption is not realistic because consumers may choose between $A$ and $B$ which implies a lower total effect (consumers have to choose). Alternatively a common price change might attract more attention and have a stronger effect.

Hence it is useful to look for other price definitions. Possible price definitions are:
The average price and the price difference

Rotated principal components. Usually these variables are similar to the average price and the price difference.

Both price definitions have serious disadvantages. First, the price difference is difficult to interpret and imposes symmetric effects. Second, the average price is not always related to a common price change. A price change for only one item also results in a change of the average price.

We solve the problem by defining separate price variables for different situations. This is similar to van Heerde et al. (2000, 2001) who define different support types for price decreases. We consider the following situations for product A’s price:

i. the price of A is not equal to the price of B
ii. the price of A is equal to the price of B

For item B we consider analogous situations:

iii. the price of B is not equal to the price of A
iv. The price of B is equal to the price of A

Note that the situations (ii) and (iv) are the same so that we have three different situations. We define three different price variables to measure the price effect in each situation. These variables are the unique price for A ($P_{UA}$) for situation (i), the unique price for B ($P_{UB}$) for situation (iii), and a common price for A and B ($P_{C}$) for situations (ii) and (iv). These price variables have the value of the price of the corresponding items if the situation occurs. An important issue is the value of these variables if the situation does not occur. We set the variables in these cases equal to the annual average price across both items. This implies that the price effect is measured as opposed to this average price. Importantly, the original prices can be computed using the new price variables, so that the new variables account for all information available in the original variables. Van Heerde et al. (2000, 2001) use a similar approach to define supported and non-supported prices. Table 5A.1. summarizes the price definitions.
Table 5 A.1 Price definitions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$P_{AI} = P_{Bi}$</th>
<th>$P_{AI} \neq P_{Bi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{UAi}$</td>
<td>$\bar{P}$</td>
<td>$P_{AI}$</td>
</tr>
<tr>
<td>$P_{UBi}$</td>
<td>$\bar{P}$</td>
<td>$P_{Bi}$</td>
</tr>
<tr>
<td>$P_{Ci}$</td>
<td>$P_{Ai}$ or $P_{Bi}$</td>
<td>$\bar{P}$</td>
</tr>
</tbody>
</table>

with $\bar{P} = \sum_{t=1}^{T} \frac{(P_{Ai} + P_{Bi})}{2}$

We illustrate the price definition with an example. Table 5A.2 shows synthetic data for two products $A$ and $B$. The table shows the observation number in the first column, the sales in the second and third columns, the prices in columns four and five, and the new price definitions in the last three columns (the average price $\bar{P}$ is 2). The prices of $A$ and $B$ are the same for the first 12 observations and different for the last 4 observations. The sales variables are generated using the following model without disturbances (we impose symmetric effects for convenience):

$$
S_{Ai} = 20 + -2P_{Ci} - 4P_{UAi} + P_{UBi}
$$

$$
S_{Bi} = 20 + -2P_{Ci} + P_{UAi} - 4P_{UBi}
$$

This model reflects the idea that a common price change of $A$ and $B$ has a smaller effect on the sales of each than a price change for one item only ($-2$ versus $-4$). Additionally the model states that the effect on the sales of $A$ and $B$ together is larger for a common price change than for a price change of one item only ($-2-2=-4$ versus $-4+1=-3$).

If we estimate model 5A.1 on this data we find own effects $\alpha_A = -3.569$ and $\beta_B = -3.436$ and cross effects $\alpha_B = 1.564$ and $\beta_A = 1.431$\(^{11}\). These effects imply that (i) an effect of a common price change on the sales of both items together is $-4.010$ ($-3.569-3.436+1.564+1.431$ ) and (ii) an effect of a price change of one item only is $-2.138$ for $A$ ($-3.569+1.431$) and $-1.872$ for $B$ ($-3.436+1.564$). It is obvious that these estimates do not reflect the difference between common and unique price changes which is non-realistic.

\(^{11}\) Estimation of model 5A.3 reproduces the model used to generate the data.
Sales decomposition using daily data for a single store

The tolerance value of the estimates indicates the percentage of variation in a predictor variable that is common with all other predictors in the model. These values are between 0 and 1. Lower tolerance values suggest that multicollinearity may be a problem (0 is perfect collinearity, 1 is no collinearity). The tolerances value for $\alpha_A$, $\alpha_B$, $\beta_A$, and $\beta_B$ in model 5A.1 are all 0.689. The tolerance value for the common price in model 5A.3 is 1.00 while the tolerance values for the unique prices are 0.06. This shows exactly the effect of the different price definitions. The tolerance values for the parameters of model 5A.1 show that the unique variation in the original price variables is limited; many changes are common. Switching to the new price definition isolates the effect of a common price change for which all variation is used for estimation (high tolerance value). At the same time there is less variation to estimate unique price effects (low tolerance values). This reflects the fact that unique price changes occur less frequently.

Table 5A.2 Price definition: an example

<table>
<thead>
<tr>
<th>Observation</th>
<th>$S_{At}$</th>
<th>$S_{Bt}$</th>
<th>$P_{At}$</th>
<th>$P_{Bt}$</th>
<th>$P_{Ct}$</th>
<th>$P_{UAt}$</th>
<th>$P_{UBt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>13</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>11</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>13</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>11</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>4.5</td>
<td>0.5</td>
<td>3</td>
<td>2</td>
<td>0.5</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>5.5</td>
<td>13</td>
<td>3</td>
<td>1.5</td>
<td>2</td>
<td>3</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Appendix 5B Coefficient transformation

For a particular day \( d \) multiplication of the estimated effects \( \beta_{ij} \) with the standard deviation \( \hat{\sigma}_{SAP,d} \) gives the additional sales effect in terms of \( SAP \) due to a change in the filtered price (FPI). This can be seen by multiplying (5.10) with the standard deviation of a particular day and rewriting the equation using (5.7) and (5.5).

\[
\hat{\sigma}_{SAP,d}SALES_{j,dw} = \hat{\sigma}_{SAP,d}C'_{j,c} + \sum_{j_c=1}^{J_c}(\hat{\sigma}_{SAP,d} \beta_{j_c,i_c} FPI_{j,dw}) + \hat{\sigma}_{SAP,d} \delta'_{j_c,i_c} FFEAT_{j,dw} + \hat{\sigma}_{SAP,d} \varepsilon_{j,dw}
\]

\( \Leftrightarrow \)

\[
\hat{\sigma}_{SAP,d}(ZSALES_{j,dw} - HP_{j,dw}) = \hat{\sigma}_{SAP,d}C'_{j,c} + \sum_{j_c=1}^{J_c}(\beta_{j_c,i_c} FPI_{j,dw}) + \delta'_{j_c,i_c} FFEAT_{j,dw} + \varepsilon_{j,dw}
\]

\( \Leftrightarrow \)

\[
SAP_{j,dw} - SAP_{j,c} = \hat{\sigma}_{SAP,d}C'_{j,c} + \hat{\sigma}_{SAP,d} HP_{j,dw} + \sum_{j_c=1}^{J_c}(\beta_{j_c,i_c} FPI_{j,dw}) + \delta'_{j_c,i_c} FFEAT_{j,dw} + \varepsilon_{j,dw}
\]

\( \Leftrightarrow \)

\[
SAP_{j,dw} = SAP_{j,c} + \hat{\sigma}_{SAP,d}C'_{j,c} + \hat{\sigma}_{SAP,d} HP_{j,dw} + \sum_{j_c=1}^{J_c}(\beta_{j_c,i_c} FPI_{j,dw}) + \delta'_{j_c,i_c} FFEAT_{j,dw} + \varepsilon_{j,dw}
\]

(5B.1)

with

- \( \beta_{j_c,i_c} \) = day specific effect of FPI (on \( SAP \)).
- \( \beta_{j_c,i_c} = \hat{\sigma}_{SAP,d} \beta_{j_c,i_c} \).
- \( \delta_{j_c,i_c} \) = day specific effect of FFEAT (on \( SAP \)).
- \( \varepsilon_{j,dw} \) = the disturbance for item \( j_c \) on day \( d \) in week \( w \) multiplied by the standard deviation of \( SAP \) on day \( d \) (\( \hat{\sigma}_{SAP,d} \)).
If we substitute the trend $T_{j,dw} = \hat{\sigma}_{j,dw} HP_{j,dw}$, and the seasonal component by $SEAS_{dj} = \hat{SAP}_{j} + \hat{\sigma}_{j,dw} C_{j}$, we obtain:

$$SAP_{j,dw} = SEAS_{dj} + T_{j,dw} + \sum_{j'=1}^{J} \left( \beta_{d_{j'}, FPI_{j,dw}} + \delta_{d_{j'}, FFEAT_{j,dw}} \right) + \epsilon_{d_{j},dw} \quad (5B.2)$$

This shows that models (5.10) and (5.13) apart from a multiplication by the standard deviation, are equivalent to a model that explains a change in sales $SAP$ from a change in filtered prices $FPI$. This sales change is measured relative to a item specific seasonal effect ($SEAS_{dj}$) and a trend $T_{j,dw}$ and is exactly the effect we want to measure (compare equation 5.4). In an analogous way the same result can be obtained for the cross category model. Note that (5B.1) cannot be estimated using OLS.

We may calculate day specific decompositions from (5.7) and (5.14). We choose to use the average effects across days multiplied by $-1$. This gives the average effect of a price decrease reported in the text:

$$-\frac{6}{d} \beta_{d_{j},j} = -\frac{6}{d} \frac{\hat{SAP}_{d_{j}}}{6} \beta_{d_{j},j} = -\frac{\hat{SAP}_{d_{j}}}{6} \beta_{d_{j},j} = \beta_{d_{j},j}^* \quad (5B.3)$$